CHILD OUTCOMES AND FAMILY SOCIO-ECONOMIC CHARACTERISTICS

FINAL REPORT OF THE PROJECT: LSAC OUTCOMES AND THE FAMILY ENVIRONMENT

BRUCE BRADBURY

SPRC Report 9/07

Social Policy Research Centre
University of New South Wales
July 2007

For a full list of SPRC Publications see, www.sprc.unsw.edu.au or contact:
Publications, SPRC, University of New South Wales, Sydney, NSW, 2052, Australia.
Telephone: +61 (2) 9385 7800 Fax: +61 (2) 9385 7838 Email: sprc@unsw.edu.au

ISSN 1446 4179

ISBN 978 0 7334 2530 1

July 2007

This research was funded by the Commonwealth Department of Families, Community Services and Indigenous Affairs (FaCSIA). The research is based upon confidentialised unit record data from the Longitudinal Study of Australian Children (LSAC). The LSAC Project was initiated and is funded by FaCSIA and is managed by the Australian Institute of Family Studies. The findings and views reported in this paper, however, are those of the author and should not be attributed to FaCSIA, the Australian Institute of Family Studies, or the Social Policy Research Centre.

Contents

Tables

Figures

Executive Summary

Which measures of family economic resources are most relevant to child outcomes? Most economic studies of child poverty in rich nations have used income as their measure of economic disadvantage. However, some have argued for broader measures, both to better reflect the actual living conditions of the disadvantaged, and to deal with the practical problems associated with income measurement.

Generally, the research addressing this measurement question has been interested in poverty or disadvantage as an outcome variable. However, the impact of economic circumstances on child outcomes is also important. This report tests the salience of different measures of family economic resources by examining the extent to which they are correlated with child learning and social/emotional outcomes (both for the disadvantaged and across the whole distribution). This is done using data from the Longitudinal Study of Australian Children (LSAC). The main focus is on the outcomes of children aged four to five years.

The outcome measures used are the summary child outcome measures for the physical, learning and social/emotional domains developed by the LSAC research consortium. A number of different indicators of family economic circumstances are used. These include

- Receipt of income support
- Income, equivalent income and full-time income (income that could be obtained if both parents worked full-time) together with low-income cut-offs (bottom 15 and 30% of families).
- Joblessness
- Subjective living-standard (how ‘getting along’)
- Hardship – whether have been unable to undertake particular activities because of shortage of money
- The ABS socio-economic indexes for areas (SEIFA) scores of the Collectors’ District (CD) or postcode in which the family lives

In some analyses we also control for other characteristics such child gender and age, family structure, mother’s age at first birth, Indigenous and non-English speaking status, and parental education.

Differences between infants and children aged 4-5

The main focus of this report is on outcomes for children aged 4-5 years (the LSAC *child cohort*). However we also assess the impact of having a low family income (or living in a low SEIFA score area) on infants, and compare this with the child cohort.

The association between family economic status and child outcomes is much stronger in the older age group. Though the outcomes for the child cohort are probably better measured than for the infants, random measurement error is unlikely to be responsible for the greater gap for the older children.

Rather, this pattern reflects a general result seen in the literature – the association of outcomes with family economic resources increases as children age. This could be due to correlations in the patterns of expression of genetic and social endowments. However, it might also reflect the cumulative impact of parental resources over time.

This second explanation is more encouraging as it suggests that environmental intervention may be able to influence child outcomes. Both explanations suggest that we should expect to find stronger correlations between family economic resources and child outcomes in future waves of the LSAC survey.

Poverty/disadvantage and child outcomes

The remainder of the report focuses on two outcome domains for the child cohort, the social/emotional outcomes and the learning outcomes.

Section 5 examines the ways in which different binary indicators of socio-economic disadvantage are associated with these outcomes. Two different thresholds are used, a 'poverty' threshold identifying approximately the most disadvantaged 15 per cent of children, and a 'disadvantage' threshold identifying around 30 per cent of children. These thresholds are arbitrary, but they are approximately the same as the fraction of children in families reliant upon income support, and the fraction of children in families receiving any income support.

The two outcome domains have different patterns of association with the economic indicators. In the case of the learning outcomes, the subjective poverty indicators (parents describing themselves as poor or very poor) have a particularly weak association. Whether the parents had approached a welfare agency is also non-significant when controlling for other variables. Most other variables are significantly associated, with the strongest correlates being joblessness and having income support as the main income source. Having three or more hardship indicators is also strongly associated, as are the income variables. (The income variables have the strongest association when the higher 'disadvantage' threshold is used). The geographic indicators have a somewhat weaker correlation with child learning outcomes (with CD level indicators only slightly better than the postcode indicators).

Contrary to the findings of other research (albeit on older children), social/emotional outcomes tend to have a stronger association with socio-economic indicators than do learning outcomes.

In the socio/emotional domain, subjective poverty has the strongest association with outcomes rather than the weakest (in contrast to the learning domain). However, because very few people actually describe themselves as poor or very poor, the variance explained by this variable remains low. The hardship, joblessness and income support receipt variables are next, followed by the income variables and then the geographic indicators. The postcode SEIFA score actually has a slightly stronger correlation than the CD level score.

Child outcomes and broader measures of parental economic resources

These associations between child outcomes and parental economic resources apply across the whole distribution rather than just to the poor and not-poor families. The adjusted R^2 statistic is used to summarise the strength of the correlation between the different resource measures and outcomes.

When we look across the whole distribution, the income variables become more powerful explanators than the subjective poverty and hardship indicators. For learning outcomes the full-time income measure has the strongest correlation. However, for social/emotional outcomes, actual income or equivalent income is more strongly associated. This reflects the fact that full-time income is determined to a considerable extent by the education level of the parents together with the fact that parental education is more strongly associated with child-learning outcomes than with child outcomes in the social/emotional domain.

On average, children from families at the 90th income percentile have an average learning score of 4.6 points higher than those at the 10th percentile. Similarly, the gap in the social/emotional score is 6 points (10 points = 1 standard deviation). When controlling for other variables (such as parental education), these associations remain significant but are much smaller (particularly for learning outcomes).

The subjective living-standard (how 'getting along') question, and the set of hardship variables, are only weakly associated with learning outcomes. However, just as for the poverty measures, these variables are one of the strongest economic correlates of social/emotional outcomes when controlling for other variables.

Though some of the specific hardship variables have stronger associations with outcomes than others, the possibility that they have the same effect cannot be rejected. This provides support for the use of a simple summary scale based on the number of hardships experienced.

For the learning outcomes, the SEIFA indicators explain less variance in the outcome variables than the income variables, but more than the subjective living-standard and hardship questions. For the social/emotional outcomes the SEIFA indicators generally explain less than all the other variables do.

Income and family size

In most poverty and inequality research, measurements of the economic welfare of households use equivalent income rather than actual family income. This takes account of the fact that larger households need more income to attain the same living standard.

In the LSAC data, learning outcomes are indeed lower when family size is larger (holding income constant). This is consistent with idea that resources matter – they will be spread thinner in the larger family. However, it might also be due to non-income related effects of family size, such as less parental time for each child. It might also be due to selection effects, e.g. having a low wage rate encourages a mother to have more children.

If we treat family size as having a pure resource-sharing effect, we can calculate an equivalence scale for learning outcomes. This shows the additional income required to offset the adverse effects of additional siblings. Compared to having only one sibling, having three siblings decreases learning outcomes so much that household income needs to more than double to offset this. This is greater than even the per-capita equivalence scale would imply. This strongly suggests that either direct or selection effects are responsible for the observed association between family size and learning outcomes.

Social/emotional outcomes, in contrast, actually improve with family size. The resource-sharing explanation doesn't account for this, so there is either a direct effect or a selection effect (or both) operating here. Having extra children in the household might aid in the development of social capacities, and/or parents who are more social might prefer to have more children.

Multiple indicators

When considered one at a time, each of the indicators has a significant association with child outcomes. Is this because these different economic indicators are highly correlated with one another, or does each of them contribute independent information on their relationship to child outcomes?

Testing for the influence of each variable while including the other indicators, it is found that most variables retain a significant association. When we also control for other variables (e.g. parental education), none of the variables has a significant impact on learning outcomes, but most do impact on social/emotional outcomes.

In the poverty measurement literature, it is common to combine different indicators to obtain a composite indicator of disadvantage. This is often done in a non-additive manner. For example, people might be described as poor if they have a low income *and* also score poorly on a hardship of disadvantage index. Do these indicators of disadvantage have a similar non-additive impact upon child outcomes?

For most pairs of variables, no significant interaction was found. This suggests that the non-additive approach is not the best approach for summarising impacts on child outcomes.

The exception is the interaction between income and SEIFA score, where the relationship between income and learning outcomes is weaker in the most advantaged regions. This could be due to the direct effect of location or, more likely, to measurement problems either in the income measures or the SEIFA measures.

Further work: developing composite indexes

Given that all the indicators examined here seem to provide some additional predictive information on child outcomes (particularly in the social/emotional domain), would it make sense to combine them into a single index of economic resources?

There are two types of composite index that could be employed – reflective or formative. Formative indexes are commonly used to create measures of socio-economic status. However, as typically used, they are essentially arbitrary in their definition. Reflective indexes (akin to factor scores) are probably more suitable in this present context. The development of these indexes would be a fruitful topic for further research.

1 Introduction

It is widely accepted that the level of family economic resources is important for child welfare – even when these resources are well above subsistence levels. However, there is continuing debate both about the mechanisms by which these resources influence child welfare, and about the best way to measure the economic resources of most relevance to children. The latter issue, which is the focus of this report, is most keenly debated in the poverty measurement literature – though many of the issues considered in this debate are equally relevant to wider measures of well-being.

Most empirical economic studies of poverty in rich nations define poverty as living in a household with a particularly low income. However, recent decades have seen a number of alternative conceptions of poverty and disadvantage advanced to challenge this conventional approach. These alternatives have been motivated partly by the perceived measurement problems associated with these conventional measures, but also by concern over the political salience of an arbitrary poverty threshold based on a purely economic measure of welfare.

Ringen (1988) has argued that indicators such as household income should be seen only as indirect indicators of poverty. To know if people are really 'going without' we should collect data directly on the extent to which they are not able to consume socially perceived necessities. Similarly, Sen's ideas of capability and functioning (1985), Townsend's (1979) relative deprivation approach, and the more recent discourse on social exclusion (eg Hills et al 2002), all seek to move away from simply using income as the yardstick of disadvantage.[1]

To these conceptual issues may be added a number of practical issues relating to the measurement of incomes. In Australia, there has been particular concern over the ability of household income surveys to accurately record the incomes of low-income families, with many families recorded as having incomes below the minimum levels of income support payments. Peter Saunders of the Centre for Independent Studies has argued that this leads to a substantial over-statement of the level of poverty in Australia (Saunders, 2004). The Australian Bureau of Statistics has cited similar concerns in its exclusion of the bottom income decile in one of its key publications (ABS, 2004). The SPRC's Saunders and Bradbury (2006) show that the bottom 3 percentiles of the income distribution do have unusually high expenditure levels,[2] and that income poverty is only weakly correlated with hardship measures. All these concerns relate to household surveys specifically designed to measure incomes. They might be expected to be even more relevant to household surveys focusing on other topics, where income information is collected using an abbreviated set of questions.

Bradshaw (2006), in reviewing variations in child poverty and child well-being across the European Union, argues that income-based child poverty measures are poor indicators of command over resources. Income is often poorly measured in household surveys, and does not include dissavings or home production. There is no consensus

1 See also Laderchi et al (2003) for a comparison of the economic, capability and social exclusion approaches.

2 Poor measurement of income among the self-employed appears to account for only a small part of this.

on how to adjust poverty thresholds across countries or between families of different size, and most income surveys do not collect information on the persistence of poverty. Bradshaw goes on to argue for the use of an index based on a number of nation-level indicators of child well-being (for the purpose of cross-national comparison). These include measures of subjective poverty (parents saying that they have difficulty or great difficulty making ends meet) and measures of deprivation (lacking a number of items from a list of deprivation indicators).[3]

In assessing the validity of these different measures of poverty, researchers have typically examined two issues: the relationship of the measures to various theoretical constructs of poverty, and the correlations between different measures. Generally, the association between low income and the alternative measures has been found to be relatively weak (e.g. Kangas & Ritakallio, 1995; Nolan & Whelan, 1996; Bradshaw & Finch, 2003; Bradshaw, 2004, Saunders and Bradbury, 2006). Researchers have also examined the correlations between the different measures, and tested the validity of generating summary indices based upon multiple indicators (e.g. Cappellari and Jenkins, 2006).

This research thus focuses on poverty as an outcome variable and examines the incidence of poverty across different groups and the causal factors associated with it. But poverty (and economic well-being more generally) is also of interest because of its anticipated effect on a wide range of other outcomes. Though we should also be concerned with poverty and deprivation as they are experienced, child poverty in particular is often given particular ethical prominence because of its likely impact upon child developmental outcomes and hence the future opportunities of the next generation.

Consequently, an alternative way to judge the salience of different measures of economic resources is to assess the association between the alternative measures and the key outcome variables that they are expected to influence. This is the objective of this study. Which of the various indicators of family economic resources are most strongly associated with the achievement outcomes of young children? Does this shed light on the best way to measure poverty (and economic resources more generally) in families with young children? Does this provide guidance for other researchers seeking to examine other factors while holding family economic resources constant?

This report addresses these questions using data from the first wave of the Longitudinal Study of Australian Children (LSAC). The focus is primarily on the developmental outcomes of the cohort of children aged 4 to 5 years at the time of interview – the *child cohort*. This report should be seen as an initial exploration of these issues. In this report, the methods used to assess these questions are relatively straightforward – primarily consisting of an examination of the correlations and partial correlations of child outcomes with different measures of economic resources. These measures include several different methods of defining family income, the socio-economic characteristics of the location in which the family lives, together with a number of measures of hardship and subjective living-standards.

[3] The measures are defined in Ritakallio and Bradshaw (2005).

The interpretation of these correlations rests upon our (limited) knowledge of the causal links between family economic resources and child outcomes. The conceptual framework for this is outlined in the following section. Section 3 of the paper then goes on to define the measures of child outcomes and family resources available in the LSAC study.

The association of income with both infant and child outcomes is shown in Section 4. Infant outcomes are only weakly associated with family income (or other socio-economic variables), and so the remainder of the paper focuses on the associations identified among the older children. Section 5 examines the link between child outcomes and indicators of poverty and extreme disadvantage, while Section 6 looks at the association with broader measures of economic resources. Section 7 looks at the relationship between income and family size, and Section 8 considers the interactions between the different economic indicators.

All the indicators of family socio-economic characteristics considered have substantial associations with child outcomes in at least one model specification. Even when controlling for the other indicators, most appear to contribute some additional information to the prediction of child outcomes – particularly in the social/emotional domain. The report thus concludes with a consideration of the possible approaches that could be used to combine these different variables into a single indicator of economic resources.

2 Economic resources and child development

Though children's outcomes are influenced by the interaction of both their genetic endowment and their environment, it is the latter that is amenable to policy intervention and is thus the focus of much research. Especially for young children, the family environment is central to their well-being. Most empirical research describing the links between the family environment and child outcomes is developed within the framework of a theoretical structure like that outlined in Figure 1.

We can divide the influences on children's developmental outcomes into proximal family processes, current family characteristics (including economic resources), and distal family characteristics. This distinction can sometimes be arbitrary, but it provides some insight into the causal links that are assumed to influence outcomes.

Proximal family processes are those most directly associated with child outcomes. They include health-related behaviours (e.g. nutrition, hygiene, accident prevention), parent-child relationships (affection and warmth, discipline), cognitive and learning-related parental behaviour (reading and explaining to children), and the quality of non-parental care.

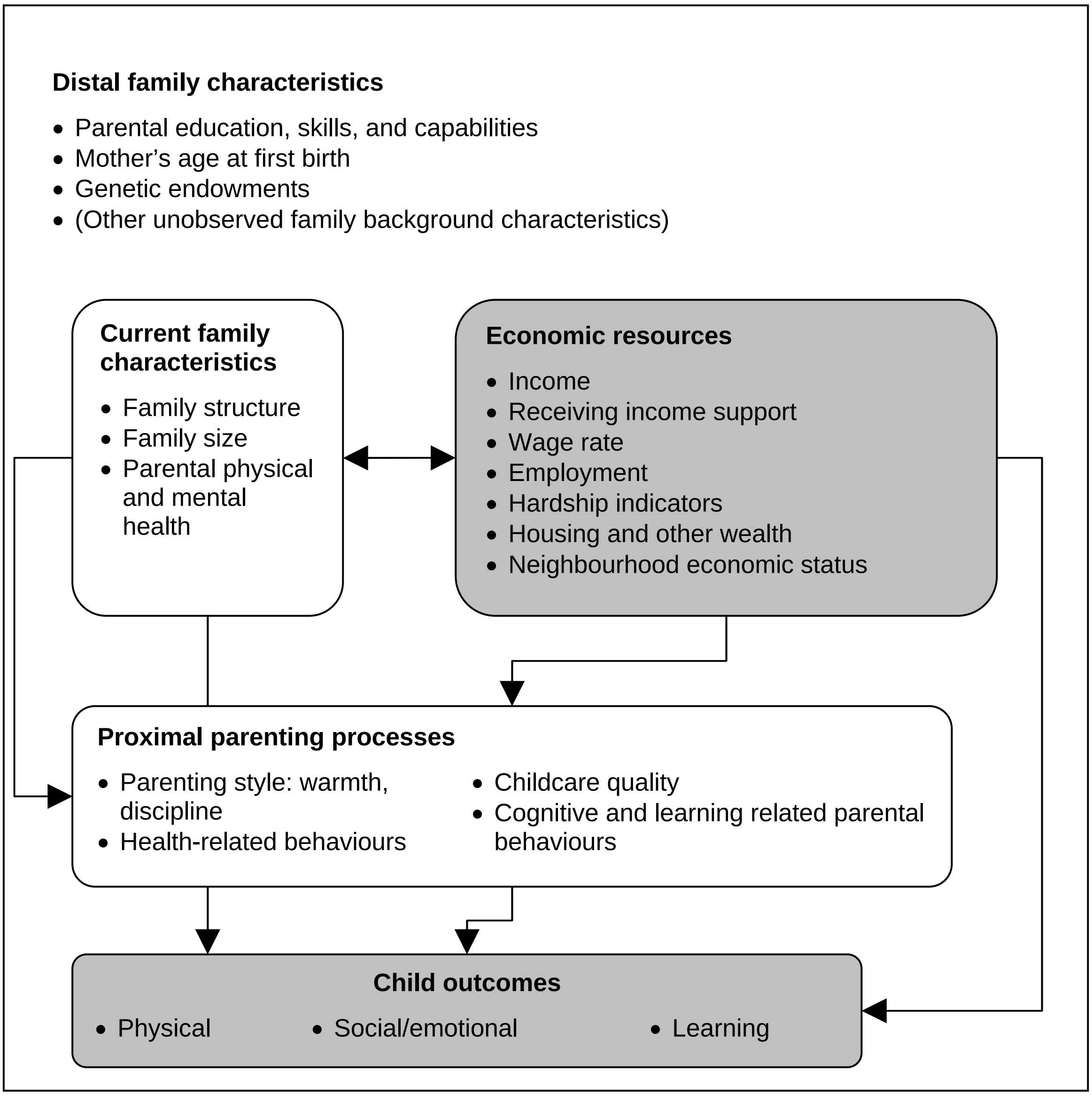
These proximal processes are in turn likely to be influenced by current family characteristics (family structure, current parental health, etc.), more long-standing family characteristics (distal characteristics such as parental capacities and indicators of social background such as age at first birth), together with the economic resources available to the family.

Economic resources also have direct impacts on child wellbeing. They permit the purchase of particular goods consumed by children such as housing, healthy food, clothing, purchased health care, childcare, and educational services. In the highly stratified cities of countries such as Australia, economic resources are an important determinant of where parents live, and hence of the characteristics of their children's peer groups and the quality of locally-delivered services.

Nonetheless, for young children, the main impact of economic resources is likely to act via their parents. For example, low income might lead to higher levels of parental stress which in turn might impair proximal parenting processes. Zubrick et al (2006) find that low-income parents are much more likely to suffer stressful life events, and to have more stressful relationships and less community connectedness. On the other hand, some factors operate in the opposite direction, with high income mothers more likely to report a lack of support and a lower degree of warmth in their parenting practices (in couples).

The focus on this report is on the overall association between the two shaded boxes in Figure 1, between different measures of family economic resources and child outcomes in the physical, social/emotional and learning spheres. To the extent to which this association is causal, some of it might be direct and some indirect via parenting processes.

Figure 1 The influence of family characteristics on child outcomes



Though observational studies such as the LSAC can capture many of the aspects of the family environment described in Figure 1, the complexity of social life means that many aspects cannot be empirically measured. In particular, there may be many unobserved distal family characteristics that influence both economic resources and child outcomes. Without controlling for these, we cannot assume that the observed association between economic resources and outcomes is causal.[4]

[4] The longitudinal data that will eventually be collected by LSAC will allow us to study the correlation of changes in family circumstances (such as a drop in income) and changes in child outcomes (such as relative achievement scores). This would hold constant any unchanging characteristics of the child or the family. However, it still would not control for unobserved characteristics that also change. For example, if the drop in income is due to an unobserved deterioration in parental mental health, we cannot tell whether any drop in child achievement is due to the income or to the mental health.

Indeed, the extent to which economic resources have a causal impact on child developmental outcomes remains controversial. Mayer (1997) argues that the causal link between family income and child outcomes is much smaller than the bivariate association. The characteristics that employers value (e.g. skills, interpersonal skills) also improve children's outcomes, independently of their effect on parental income. Moreover, the usual range of socio-economic controls available in household surveys cannot capture all these characteristics.

One way of addressing this question is to compare siblings who share the same parents, but who were raised when family incomes were at different levels. Levy and Duncan (2000) use this approach and conclude that family income does have an effect on child outcomes, though it still remains much smaller than the raw association would suggest.[5]

More recent research has drawn on the independent variation introduced by the welfare-to-work experiments conducted in the US in the early 1990s and the subsequent policy changes. Morris, Gennetian and Duncan (2005), review the results from these experiments and find improvements in school learning outcomes (albeit relatively small) when young children have parents who participate in welfare programs. By comparing experiments with different mixes of employment and income increases, they conclude that these effects are due to the income increases (and possibly increased use of centre-based childcare). Dahl and Lochner (2005) compare families facing different policy changes under the Earned Income Tax Credit (EITC) of the last two decades, and also find a small but significant effect of income ($1000 per annum income increase associated with 2% and 3.6% increase in math and reading test scores respectively).

Here we proceed by assuming that some of the association between parental economic resources and child outcomes is probably causal, but much of it simply reflects unobserved characteristics that influence both parental economic outcomes and child developmental outcomes. In presenting the results below, causal language ('impact' 'effect') is sometimes used for convenience. But in examining which measure of economic resources is most strongly associated with child outcomes, we are not attempting to ascribe causality, but rather simply to say which measure best summarises the impact on children of this combination of economic resources and other correlated factors.

Note also that the focus here is on measures of economic resources. Other studies, such as Blakemore, Gibbings and Strazdins (2006, following Willms and Sheilds, 1996), have adopted a broader sociological approach and included measures of educational attainment and the status of the parents' occupations in their socio-economic status measures. We adopt a narrower focus here primarily for reasons of conceptual clarity. Characteristics such as education level and status might affect parenting directly in ways that are additional to their impact on access to the goods and services of a market economy. (The 'full-time income' measure developed below incorporates information on the association between education, occupation and income).

5 See also Duncan et al (1998).

3 Data and variable definitions

3.1 The LSAC

The Longitudinal Study of Australian Children (LSAC) (also known as 'Growing up in Australia') is a longitudinal study of two cohorts of children.[6] This report uses data from the first wave of the study. The first cohort, the *infant cohort*, is represented by a sample of around 5,100 infants who were born between March 2003 and February 2004. Their families were first interviewed when they between 3 and 19 months old. The second cohort, the *child cohort*, comprises 5,000 children born between March 1999 and February 2000. They were between 51 and 67 months (4¼ to 5½ years old) when their family was interviewed. Most of the results here (except for Section 4) are for the older cohort only.

The sample was selected from the Medicare database, with participants able to opt-out before approached by the researchers.[7] The main data collection was a face-to-face interview with the person who knew the child best, usually the mother. She provided information about herself, the child and her co-resident spouse if relevant. Some direct assessments of the child were also made by the interviewer, and self-completion questionnaires were left for the two parents. The response rate was 57 and 50 per cent of those people initially selected for the infant and child cohorts respectively.

3.2 Child outcome measures

The LSAC survey collects a wide range of information on child developmental outcomes. This report draws on the summary outcome measures developed by Sanson, Misson et al (2005). These have been developed for both the infant and child cohorts, though they are considered more reliable for the latter group – which is thus the focus of this study.

Sanson, Misson et al define summary indicators for three separate sub-domains, Physical, Social/Emotional and Learning (we focus mainly on the latter two domains here). The scores in each domain are standardised to have a mean of 100 and a standard deviation of 10. Higher scores indicate better functioning, though it should be noted that the items used to generate the summary measures are more relevant to problems than to capacities at the high end of achievement. The items used to generate the sub-indices for the child cohort are summarised in Table 1. The teacher-rated scores were generally only available for about two-thirds of the sample. In these cases, the child's score is based on the remaining measures. See Sanson et al Appendix A for the methods used.

6 See Australian Institute of Family Studies (2006) for details of the study design.

7 The sample was clustered by postcode and stratified by region, with some remote areas excluded.

Table 1 Components of the LSAC outcome measures for children aged 4-5

Physical	
Overall health rating	Single parent-rated item of child's health
Special health care needs	Single item indicating whether child needed medication or more health care than the average child due to a condition that has lasted or was expected to last 12 months or more
Body-mass index	Directly measured height and weight
PEDS QL Physical health subscale	8-item parent report (motor coordination and general health)
Social/Emotional	
SDQ Prosocial	5 parent-rated items assessing the child's propensity to behave in a way that is considerate and helpful to others
SDQ Peer problems	5 parent-rated items assessing problems in the child's ability to form positive relationships with other children
SDQ Emotional	5 parent-rated items assessing a child's frequency of display of negative emotional states (e.g. nervousness, worry)
SDQ Hyperactivity	5 parent-rated items assessing child's fidgetiness, concentration span and impulsiveness
SDQ Conduct	5 parent-rated items assessing child's tendency to display problem behaviours when interacting with others
Learning	
PPVT	Interviewer administration of an abbreviated Peabody Picture Vocabulary test
Parent rating of reading skills	3 items assessing whether a child has obtained reading skills at different levels of complexity
Teacher rating of reading skills	5 items assessing the level of complexity a child is capable of reading and the child's interest in reading.
Teacher rating of writing skills	6 items assessing the level of complexity of the child's writing skills as well as the child's interest in writing.
Teacher rating of numeracy skills	5 items assessing the child's ability to perform numeric tasks such as counting, classifying, and simple addition, along with the ability to recognise numbers
Who Am I	Interviewer administration of a measure which assesses a child's ability to perform a range of tasks such as reading, writing, copying, and symbol recognition, as a measure of school readiness.

Source: Sanson, Misson et al (2005).

3.3 Measures of family economic resources

Several different measures of family economic resources are collected in the LSAC survey. Some measures apply to all families, while others are more relevant to the measurement of particularly low levels of economic resources. The measures used in this report are summarised below.

Income-based measures

Income support is main income source (1=yes, 0=no). This is defined as either mother or father having an income support pension or benefit as their main income source and neither having anything else as their main income source. Approximately 16 per cent of the families of the 4 to 5-year-old cohort have income support as their main income source.

Any income support received (1=yes, 0=no). Mother or father is receiving Parenting Payment Partnered, Parenting Payment Single, Newstart Allowance or Disability Support Pension.

Income. The respondent (P1 in the survey terminology, usually the mother) is asked about her own usual income from all sources before tax is taken out (see the Appendix for question wording). She is also asked about her spouse's income (P2). She is also asked about the present yearly income for herself and her partner combined (in categories). For most families, income is defined as the sum of the respondent and spouse income. If there is no spouse present, his income is set to zero for this calculation. If this sum is missing, then the mid-point of the categorical response is used. All incomes are converted to weekly equivalents.

Note that this is a much abbreviated way of measuring family income, compared to that used in dedicated income surveys (such as the ABS Income surveys and the HILDA survey). These surveys typically ask a long series of questions, seeking to separately identify the amount of income received from each income source. Abbreviated questions such as that used here are likely to lead to underestimation of income because people forget about income from minor income sources such as family payments. On the other hand, the LSAC abbreviated question is typical of that used in many other household surveys which have non-economic topics as their main focus.

Bottom 15% of income / bottom 30% of income. Family has incomes in the bottom 15 or 30 per cent of the (weighted) income distribution. The 15 and 30 per cent thresholds are arbitrary, but are chosen to approximate the percentages of the population with income support as their main income source, or receiving any income support.

Equivalent income (and bottom 15 and 30% of this). This is income divided by the square root of the number of people in the household. This is a crude way of taking account of the additional needs of larger families. It does not distinguish between adults and children (though many of the results here control for family structure, which has the same effect). It is also approximate to the extent that only the income of the mother and her spouse are recorded, but household size might depend upon other adults. However, this is likely to apply only to a small fraction of the sample.

Full-time income (and bottom 15 and 30% of this). The labour-force participation of the mothers of young children varies substantially depending upon the childcare arrangements made for their children. To a certain extent, therefore, family income will depend upon the choices made by the parents, rather than the opportunities available to them. In order to focus more clearly on the latter, we calculate the income that the family could have received if both the mother and father (if present) were working full-time. Where the mother or father is not working full-time, we impute their full-time income based upon their educational attainment, age, number of children (an indicator of labour-market experience), and their last occupation. See the Appendix for the imputation details. Where the parents are working full-time we use their actual income.

Social exclusion

Though social exclusion has been measured in many ways, one core feature that always appears is a lack of engagement with the labour market. A variable indicating that *neither parent has a job* is thus used as a social exclusion indicator. In 13 per cent of families neither parent had a job.

Lack of employment clearly reduces family income, but it might also have additional negative impacts upon family life. For older children these might include role model effects, but even for the young children in this study, the lack of structure associated with joblessness might lead to parental stress and poorer parenting practices.

Subjective well-being

Survey respondents' own subjective views on their economic circumstances are particularly useful when their needs vary in ways that are not known to researchers. In the LSAC survey, respondents were asked "Given your current needs and financial resources, how would you say you and your family are getting along?" with response options of: prosperous, very comfortable, reasonably comfortable, just getting along, poor or very poor.

We use binary dummy variables for all but one of these categories (*how getting on*) and also (separately) consider two disadvantage indicators for those reporting *just getting along, poor or very poor*, and for those reporting *poor or very poor*.

Hardship

A more concrete way of assessing disadvantage, and more in keeping with Stein Ringen's concept of direct poverty, is to collect information on particular activities that people have undertaken because of their low level of economic resources (or possibly because of their high needs).

The LSAC survey asks the primary respondent whether:

> Over the last 12 months, due to shortage of money, have any of the following happened?
>
> - You have not been able to pay gas, electricity or telephone bills on time?
> - You could not pay the mortgage or rent on time?
> - Adults or children have gone without meals?
> - You have been unable to heat or cool the home?
> - You have pawned or sold something?
> - You had financial limits on the type of food you could buy?
> - You have sought assistance from a welfare or community organisation?

We include binary variables for each of these events (*shortage of money events*) as well as a number of summary indicators.[8] In terms of easily observable characteristics, many might consider seeking assistance from a welfare organisation a strong indicator of need, and so *seeking assistance from a welfare agency* is one indicator.

8 For all these binary variables, the small number of missing cases are treated as 'no'.

Saunders and Bradbury (2006), however, argue that this indicator (together with the *pawned or sold something* question) might reflect a conscious decision that might be influenced by lack of information, shame or stigma. Their summary indicators (from a similar set of questions in the 1998-99 ABS Household Expenditure Survey) exclude questions of this type. Following this approach we also define two poverty indicators *two or more hardship indicators* and *three or more hardship indicators* where the indicators are the five items from the list above, but excluding the welfare and pawned questions.

The LSAC survey also includes a question on whether the household has *private health insurance*. We include this also as an indicator of advantage rather than poverty, though it is possible that it also reflects parents' attitudes to health care.

Locality-based indicators

In Australia, it is increasingly common for researchers to use measures of the characteristics of the neighbourhood or suburb as a measure of socio-economic resources. At the practical level, this approach is prompted by the difficulties of measuring reliable income data at the household level and by the easy availability of the ABS-created socio-economic indexes for areas (SEIFA). More generally, these measures are probably strongly correlated with land prices (not easily available to researchers in Australia) and might thus be a good indicator of total wealth and household permanent income levels.

In addition, community characteristics themselves might have a direct impact on children's outcomes. In more socially advantaged areas, the quality of services such as health and childcare might be higher, and the other children attending these services will come from more advantaged families.

The ABS produces four SEIFA indexes from each Census. They are based on the average characteristics of the people living in the locality, using Census-collected variables such as income, education and qualifications, etc. Higher scores indicate more advantaged characteristics. See Adhikari (2006) for an introduction.

The index of relative socio-economic disadvantage (IRSD), is based on indicators such as low income, unemployment, and low levels of education. The index of relative socio-economic advantage/disadvantage (IRSAD) is conceptually similar, except that it also includes indicators of advantage (such as the percentage of the workforce in professional employment, high income levels, etc.). The index of economic resources focuses on income and housing characteristics, while a fourth index focuses on education and occupation. All four of the SEIFA indices have been used here (though when we look at poverty and disadvantage in Section 5 we focus on the index of disadvantage).

For the 2001 Census all these indicators these were calculated at the Collectors' District (CD) level (about 200 households on average). Estimates for larger areas were calculated by aggregation. This aggregation could potentially lose some information. For example, a particular statistical local area (SLA) (the same as a local government area in some States) might contain both rich and poor neighbourhoods. Adhikari (2006) examines the correlation of the IRSD index with health status and shows that the correlation is much stronger when the CD rather than the SLA level indexes are used.

For the LSAC survey, only the postcode-level SEIFA indexes are provided in the public use file in order to protect respondent confidentiality. Though postcodes are generally smaller than SLAs, it is still of interest to see if this aggregation significantly reduces the information content of the indexes. Consequently we also examine the associations with CD-level indicators (specially calculated by AIFS for this project).

3.4 Control variables

We are interested here both in the raw association between the indicators of family economic resources and child outcomes, and also in the partial relationship, controlling for other factors that might influence child outcomes. Some of these factors could be expected to be orthogonal to economic resources,[9] but they are included to increase the precision of the estimates. These variables include the child's gender and their age in months.

Other distal and current variables are known to be associated with both outcomes and economic characteristics. These include: whether the child is in a lone parent household, whether one of the parents present is non-biological, whether their mother is not legally married, the number of children in the family, whether the child is the first-born, the mother's age when she had her first child (and squared), whether the mother ever smoked regularly, whether the child is Indigenous, whether the main language spoken by the child is not English, and mother's and father's schooling attainment and post-school qualifications (sets of dummy variables).

Some of these variables, particularly education, are often included as components of socio-economic status in their own right (e.g. Blakemore et al, 2006). We don't follow this approach as we seek to focus on the economic resources available to families rather than on the broader set of cultural resources associated with income. The full-time income measure described above does, however, take account of education to the extent that it is correlated with income (among full-time workers).

[9] Variables such as the child's age when interviewed are unlikely to be systematically correlated with economic status. Hence including them in a regression will not change the estimated impact of economic status, except in so far as there is a random association between the two variables in this particular sample.

4 The different impact of family income on infant and child outcomes

The main focus of this study is on the associations between indicators of family economic resources and the outcomes of the child cohort. Before examining these indicators, however, it is instructive to compare the patterns for the infants with the older child cohorts. A striking fact that emerges from this comparison is that the association of outcomes with income is much stronger for the older children.

Table 2 shows the association between low family income (bottom 15% of the income distribution[10]) and outcomes in the physical, social-emotional and learning domains for infants and for children. Both the raw associations and the associations involving the control variables described above are shown (estimated by OLS regression in both cases). The top-left cell, for example, indicates that children in the bottom 15 per cent of family income have a 0.64 lower score on the physical domain summary index. The summary indexes all have means of 100 and standard deviations of 10, so this implies an effect size of 0.06 of a standard deviation on physical outcomes. The approximate t statistic implies that this association is not significantly different from zero. Controlling for other variables, the association is negligible (and in the opposite direction).

For the children, the effect size for the simple association is about twice as large.

Table 2 Impact of being in the bottom 15 per cent of family income on infant and child outcomes

	Simple association		Controlling for other variables	
Model	Estimate	approx t	Estimate	approx t
Physical domain				
Infants	-0.64	-1.4	0.07	0.1
4-5 year olds	-1.22	-2.8	-0.06	-0.1
Social-emotional domain				
Infants	-0.80	-1.6	-0.17	-0.3
4-5 year olds	-4.01	-9.1	-1.48	-2.8
Learning domain				
Infants	2.12	4.5	0.59	1.0
4-5 year olds	-3.49	-7.9	-1.27	-2.5

Source: LSAC survey, wave 1. The control variables are: sex, age (months), Indigenous, non-English speaker, 1st born, age youngest child, n children, n persons, family type, mother's age at first birth, mother smoked, and parents' educational attainments. Absolute values of the t statistic greater than 1.96 indicate the estimate is significantly different from zero at the 5% level.

For the social-emotional summary score, the difference between the infants and the children is even larger. For a kindergarten-age child, being in the bottom 15 per cent of the income distribution means having a social-emotional outcome 0.4 standard deviations lower (-4.01/10). This continues to be statistically significant even when controlling for the other family characteristics. (Recall that the controls include strong socio-economic predictors such as parental education and mother's age at first birth).

10 The patterns are very similar for equivalent income.

Surprisingly, the learning outcomes for infants are associated in the opposite direction to that expected. Low-income infants tend to have significantly *higher* scores on the learning domain measures. For the children this association reverses to the expected pattern, though not quite as strong as for the social-emotional domain. Again the relationship remains significant when controlling for other variables.

Table 3 shows that similar patterns occur when we use an alternative measure of family economic resources. In this table, disadvantaged families are identified by whether they fall into the 15 per cent of families in the survey living in the areas with the lowest scores on the SEIFA Index of Disadvantage. Using this indicator, the impact of disadvantage is always negative. Nonetheless, the raw effect size is always larger for the older children. Why should this be so?

Table 3 Impact of being in the bottom 15 per cent of the SEIFA Index of Disadvantage on infant and child outcomes

	Simple association		Controlling for other variables	
Model	Estimate	approx t	Estimate	approx t
Physical domain				
Infants	-1.25	-2.9	-0.88	-2.0
4-5 year olds	-1.42	-3.3	-0.88	-2.0
Social-emotional domain				
Infants	-1.36	-2.9	-0.94	-1.9
4-5 year olds	-3.50	-8.2	-1.76	-4.2
Learning domain				
Infants	-0.06	-0.1	-0.99	-2.1
4-5 year olds	-2.64	-6.2	-0.76	-1.9

Source: LSAC survey, wave 1. The control variables are: sex, age (months), Indigenous, non-English speaker, 1st born, age youngest child, n children, n persons, family type, mother's age at first birth, mother smoked, and parents' educational attainments. Absolute values of the t statistic greater than 1.96 indicate the estimate is significantly different from zero at the 5% level.

One potential explanation is that the domain indexes are, by necessity, different for each age group. In particular, the LSAC collects much more detail on the outcomes for the older children – outcomes that could not sensibly be measured for infants. For the infants, the outcomes scores are based on a range of parent-rated items. For the older group, these are supplemented by the BMI index for the physical domain, and interviewer-administered tests and teacher ratings for the learning domain. In addition, all three domains ask additional questions of the parents of the older children. (See Sanson, Misson et al 2005 for the details of the indexes).

It is thus likely that the developmental outcomes of the younger group are more poorly measured than those of the older group. Indeed, it may be the case that it is intrinsically impossible to measure developmental outcomes of infants with the same accuracy as for older children. However such measurement error, of itself, should not lead to the patterns of association found here. Random measurement error in the dependent variable in a regression will reduce the precision of the estimates (and the R^2 and t statistics), but will not bias the estimates of the effect of independent variables (low income in this case). Measurement error in the predictor variables (income in this case) will bias the results, but there is no reason to believe that income is measured more poorly among the parents of infants than among the parents of children aged 4-5.

Nonetheless, systematic (as opposed to random) differences in the measurement methods might lead to these results. In the learning domain in particular, the older-cohort score is much more heavily based than is the infant cohort score, on external measurements rather than the mother's own ratings of the child's abilities. It is possible that the lower expectations of low-income mothers might lead them to rate their child's activities more highly on the questionnaire items. However, the same between-cohort differences also exist in the social-emotional domain, which is also based on parent ratings in both cohorts.

Moreover, this increase in the correlation between economic status and child outcomes over the pre- and early school years has been found by previous researchers using other data collections. Feinstein (2003) reports on the early cognitive development of British children, and finds a steady widening of the test score gap between the ages of two and five, between children from high SES families (grouped by father's occupation) and those from low SES families. In the US, Fryer and Levitt (2004) examine the test score gap between black and white children over the first four years of school. They also find a steadily widening gap that cannot be explained by observable characteristics (including school quality).[11] Cunha et al (2005) emphasise the substantial socio-economic gap in mathematics skills that exists by age six, but also show a widening of this gap over the subsequent six years.

There are two types of explanation for this pattern of widening socio-economic gaps. On the one hand, they might simply be a reflection of the patterns of correlation between genetic capabilities of parents and their offspring. Childhood is a time of development, with different capabilities arising at different times. Those capabilities that arise later are likely to be more strongly associated with adult outcomes such as income. Hence we should expect to find a stronger correlation between parental and child achievements as the children get older.

On the other hand, this pattern might reflect the cumulative impact of parental resources (or environmental factors correlated with them) on child outcomes. This explanation is more encouraging as it suggests that some form of environmental intervention might be able to influence child outcomes.

For the reasons discussed in the previous section, it is very difficult to empirically separate these two classes of explanation. Nonetheless, both explanations suggest that the significant associations found here between child outcomes and family economic circumstances will grow stronger in future waves of the LSAC survey.

[11] See also Andrew Leigh's presentation at the 2006 FaCSIA SPRS conference for similar evidence for Aboriginal and non-Aboriginal Australians.

5 Which measures of parental poverty and disadvantage have the strongest associations with outcomes for the child cohort?

In the remainder of this report we focus on outcomes for the child cohort in the social-emotional and learning domains. (As shown above, economic associations with physical outcomes are weaker). This section examines the impact of poverty and disadvantage on these outcomes. Which indicators of disadvantage are most strongly associated with child outcomes?

The indicators of disadvantage used are shown in Table 4. They are grouped here under two labels 'poverty' and 'disadvantage'. These terms do not signal any qualitative difference between the two sets of indicators, but simply that the first group refers to those families with the lowest scores on the resource indicator, while the second group includes those who have slightly higher scores. The poverty indicators mainly apply to just under 15 per cent of the sample, and the disadvantage indicators to around 30 per cent. These thresholds were chosen for convenience as round figures, respectively similar to the proportions of families with income support as their main income source or with any income support received.

Note that 13 per cent of children live in families with incomes in the bottom 15 per cent of the population. This discordance arises from the fact that this table is for those cases that had valid values on all the key indicator variables. Also, the means shown in this table are unweighted while the 15 per cent threshold is based on weighted data.

Table 4 Means of poverty and disadvantage indicators

Economic status variables	Unweighted mean
Poverty indicators (<15% of sample)	
Income support is main income source	0.13
Bottom 15% of income	0.13
Bottom 15% of equivalent income	0.12
Bottom 15% of full-time income	0.14
Neither parent has job	0.09
Poor or very poor	0.03
Has approached welfare agency	0.06
Three or more hardship indicators	0.08
Bottom 15% of CD SEIFA disadvantage	0.13
Bottom 15% of Postcode SEIFA disadvantage	0.13
Disadvantage indicators (25-35% of sample)	
Any income support is received	0.27
Bottom 30% of income	0.26
Bottom 30% of equivalent income	0.26
Bottom 30% of full-time income	0.29
Just getting along, poor or very poor	0.36
Two or more hardship indicators	0.22
Bottom 30% of CD SEIFA disadvantage	0.30
Bottom 30% of Postcode SEIFA disadvantage	0.30
Advantage indicator	
Has private health insurance	0.48

Source: LSAC child cohort, wave 1. Cases missing on any variable excluded. N = 4,412.

Among the poverty indicators, most identify around 13 per cent of the sample. Exceptions are *neither parent has a job* (9%), self-expressed poverty (3%), *has approached welfare agency* (6%) and *three or more hardship indicators* (8%). Among the disadvantage indicators, most have an incidence of around 30 per cent. The outliers are *just getting along, poor or very poor* (36%) and *two or more hardship indicators* (22%).

Table 5 shows the relationship between the child cohort learning domain scores and the different poverty and disadvantage indicators. The first line of the table shows the adjusted R^2 of an OLS regression model using all the control variables described in the previous section (31 variables). These explain about 15 per cent of the variance in the learning domain score across children. The second panel shows the estimates from a series of one-variable regressions of each economic variable in turn. Most poverty or deprivation indicators explain about 1.5 to 2.5 per cent of the variance. When they are added to models already including the control variables, they explain only a very small amount of additional variance (i.e. compare the adjusted R^2 in the 'with control variables' panels of the table with that for the control variables alone). Though comparisons of adjusted R^2 measures are meaningful, their absolute values are not particularly informative as they will be influenced by the precision of measurement in the dependent variable (among other things).

For these poverty and deprivation indicators the most meaningful indicator is the effect size shown in the parameter estimate column. This shows the impact of poverty or disadvantage on learning outcome scores. For the poverty indicators without any control variable, these range from –2.2 to –4.8. Since the outcome scores have a standard deviation of 10, a parameter estimate of 5 corresponds to an effect size of half a standard deviation – a noticeable but not extremely large impact. By way of comparison, the parameter associated with being female is around 4.2, as is the impact of being 10 months older.

Considering the poverty indicators on their own (first panel) the economic indicator with the strongest links with child learning outcomes is the joblessness of the parents (–4.8). Income support as the main income source is next, followed by experiencing three or more hardship indicators (–3.9). Having approached a welfare agency and the three different income measures have slightly smaller impacts (–3.5 to –3.8).

Equivalent income has a stronger association than actual income – implying that living in a larger family is associated with a lower learning score. We consider this relationship further in Section 7. Interestingly, full-time income does not have any stronger association with learning outcomes than does income on its own. This could be because the imputation process introduces additional measurement error into this variable (i.e. the value is the full-time income imputed for the person, not the actual income they would have received). To estimate the true impact of full-time income on outcomes, one should control for this attenuation. However, here we are interested in the best empirical predictor of outcomes for which we must include any attenuation due to measurement error.

The geographic indicators have a somewhat weaker association with outcomes, with the index created at the CD level higher (but not dramatically so) than the postcode indicator. Finally, the subjective poverty indicator (people considering that they are poor or very poor) has the weakest association with child outcomes (though only 3% of children fell into this category).

Table 5 Poverty and disadvantage indicators and learning domain scores

Model	Adj. R^2 (%)	Parameter estimate	approx t	Parameter estimate rank
Control variables	15.1			
	Poverty (low threshold)			
Poverty indicators alone				
Income support is main income source	2.0	-4.2	-9.6	2
Bottom 15% of income	1.4	-3.5	-7.9	7
Bottom 15% of equivalent income	1.6	-3.8	-8.5	4
Bottom 15% of full-time income	1.6	-3.5	-8.5	6
Neither parent has job	2.1	-4.8	-9.8	1
Poor or very poor	0.1	-2.2	-2.6	10
Has approached welfare agency	0.7	-3.6	-5.7	5
Three or more hardship indicators	1.2	-3.9	-7.2	3
Bottom 15% of CD SEIFA disadvantage	1.1	-3.0	-7.0	8
Bottom 15% of Postcode SEIFA disadvantage	0.8	-2.6	-6.2	9
Poverty indicators with control variables				
Income support is main income source	15.3	-1.5	-2.7	2
Bottom 15% of income	15.3	-1.3	-2.5	5
Bottom 15% of equivalent income	15.2	-0.9	-1.8	7
Bottom 15% of full-time income	15.2	-1.3	-1.5	4
Neither parent has job	15.4	-2.0	-3.5	1
Poor or very poor	15.1	-0.3	-0.3	10
Has approached welfare agency	15.1	-0.6	-1.0	9
Three or more hardship indicators	15.3	-1.5	-2.8	3
Bottom 15% of CD SEIFA disadvantage	15.2	-0.9	-2.2	6
Bottom 15% of Postcode SEIFA disadvantage	15.2	-0.8	-1.9	8
	Disadvantage (high threshold)			
Disadvantage indicators alone				
Any income support is received	1.9	-3.0	-9.2	4
Bottom 30% of income	2.2	-3.3	-9.9	3
Bottom 30% of equivalent income	2.6	-3.6	-11.0	1
Bottom 30% of full-time income	2.5	-3.4	-10.8	2
Just getting along, poor or very poor	1.0	-2.1	-6.8	8
Two or more hardship indicators	1.3	-2.7	-7.8	5
Bottom 30% of CD SEIFA disadvantage	1.3	-2.4	-7.6	6
Bottom 30% of Postcode SEIFA disadvantage	1.0	-2.2	-6.9	7
Disadvantage indicators with control variables				
Any income support is received	15.2	-0.6	-1.5	7
Bottom 30% of income	15.3	-1.1	-2.9	2
Bottom 30% of equivalent income	15.3	-1.2	-3.3	1
Bottom 30% of full-time income	15.3	-1.0	-2.5	3
Just getting along, poor or very poor	15.2	-0.5	-1.8	8
Two or more hardship indicators	15.2	-0.6	-1.8	4
Bottom 30% of CD SEIFA disadvantage	15.2	-0.6	-2.0	5
Bottom 30% of Postcode SEIFA disadvantage	15.2	-0.6	-2.0	6
	Advantage indicator			
Has private health insurance	3.5	3.7	12.7	
Advantage indicator with controls				
Has private health insurance	15.4	1.1	3.5	

Source: LSAC child cohort, wave 1. Cases missing on any variable excluded. N = 4,412. See Table 2 for control variables. Absolute values of the t statistic greater than 1.96 indicate the estimate is significantly different from zero at the 5% level.

The second panel of the table examines the impact of these variables when holding the other characteristics of the family constant. The effect size drops dramatically because these variables are strongly correlated with characteristics such as parental education, family structure and mother's age at first birth. Nonetheless, the ranking of the effect size is very similar to that in the first panel, and most variables remain statistically significant.[12] The main change is that having approached a welfare agency is now the second weakest correlate (and is not significant).

The bottom half of the table shows a similar analysis undertaken for a range of indicators based on a higher disadvantage threshold. Here the strongest indicators are the three income measures (equivalent income the strongest). The subjective poverty measure is again the weakest, with the SEIFA indicators slightly stronger. When controlling for other variables, the income-support-received variable is no longer significant.

Finally, the one indicator of advantage, having private health insurance, is moderately associated with child learning outcomes. Because it applies to about half the population, it also explains a relatively large fraction of the variance.

To summarise these relationships between learning outcomes and family economic indicators: Non-employment and having income support as main income source are most strongly associated child outcomes. Having three or more hardship indicators is also strong, as are the income variables. (The income variables are strongest when a higher threshold is used). The geographic indicators are weaker, with subjective poverty having the weakest relationship with child-learning outcomes.

Table 6 shows the corresponding analysis for child social/emotional outcomes. The results are quite different from the learning outcomes in many respects. First, the amount of variance explained by the control variables is much less (9% vs 15%). This could be due either to a weaker association between family characteristics and child social/emotional outcomes, or to more measurement error in the social/emotional domain scores.

As noted above, the impact of being female is now about half of that in the learning domain, and the effect of age is now negligible.[13] However, the association with economic variables is generally stronger than for the learning domain. This latter result is in contrast with the 'stylised fact' reported by Duncan et al (1998), that family income has a stronger association with achievement scores than with behavioural outcomes. Duncan et al are referring to later-life outcomes, however, which might be only loosely associated with social/emotional scores at age 4-5.

12 The t statistics are based on random sampling formula here. Because of clustering, the true t values will be slightly lower than those shown.

13 For the learning domain, the effect size of being female is around 4.2, as is the impact of being 10 months older. For the social/emotional domain, these effect sizes are 2.3 and 0.2 respectively.

Table 6 Poverty/ disadvantage and social/emotional domain scores

Model	Adj. R^2 (%)	Parameter estimate	approx t	Parameter estimate rank
Control variables (distal and current)	8.9			
Poveryt (low threshold)				
Poverty indicators alone				
Income support is main income source	2.2	-4.4	-10.1	5
Bottom 15% of income	1.8	-4.0	-9.1	6
Bottom 15% of equivalent income	1.3	-3.5	-7.7	9
Bottom 15% of full-time income	1.6	-3.6	-8.6	7
Neither parent has job	2.3	-5.1	-10.2	3
Poor or very poor	1.0	-5.6	-6.7	1
Has approached welfare agency	1.4	-4.9	-7.9	4
Three or more hardship indicators	2.0	-5.1	-9.6	2
Bottom 15% of CD SEIFA disadvantage	1.3	-3.3	-7.8	10
Bottom 15% of Postcode SEIFA disadvantage	1.5	-3.5	-8.2	8
Poverty indicators with control variables				
Income support is main income source	9.1	-1.5	-2.7	6
Bottom 15% of income	9.1	-1.5	-2.8	7
Bottom 15% of equivalent income	8.9	-0.5	-1.1	10
Bottom 15% of full-time income	8.9	-0.6	-0.6	9
Neither parent has job	9.2	-2.2	-3.6	4
Poor or very poor	9.4	-4.0	-4.9	2
Has approached welfare agency	9.3	-2.6	-4.0	3
Three or more hardship indicators	2.0	-5.1	-9.6	1
Bottom 15% of CD SEIFA disadvantage	9.1	-1.4	-3.3	8
Bottom 15% of Postcode SEIFA disadvantage	9.3	-1.8	-4.2	5
Disadvantage (high threshold)				
Disadvantage indicators alone				
Any income support is received	3.0	-3.8	-11.7	3
Bottom 30% of income	3.2	-4.0	-12.1	2
Bottom 30% of equivalent income	2.7	-3.7	-11.1	4
Bottom 30% of full-time income	2.3	-3.3	-10.3	5
Just getting along, poor or very poor	2.4	-3.1	-10.4	6
Two or more hardship indicators	3.2	-4.2	-12.1	1
Bottom 30% of CD SEIFA disadvantage	1.3	-2.5	-7.8	8
Bottom 30% of Postcode SEIFA disadvantage	1.9	-2.9	-9.3	7
Disadvantage indicators with control variables				
Any income support is received	9.4	-1.9	-5.0	2
Bottom 30% of income	9.4	-1.9	-4.7	3
Bottom 30% of equivalent income	9.2	-1.4	-3.7	6
Bottom 30% of full-time income	9.0	-1.0	-2.5	8
Just getting along, poor or very poor	9.7	-1.9	-6.1	4
Two or more hardship indicators	10.1	-2.7	-7.5	1
Bottom 30% of CD SEIFA disadvantage	9.1	-1.1	-3.3	7
Bottom 30% of Postcode SEIFA disadvantage	9.5	-1.6	-5.2	5
Advantage indicator				
Has private health insurance	3.5	3.6	12.6	
Advantage indicator with controls				
Has private health insurance	9.4	1.6	5.0	

Source: LSAC child cohort, wave 1. Cases missing on any variable excluded. N = 4,412. See Table 2 for control variables. Absolute values of the t statistic greater than 1.96 indicate the estimate is significantly different from zero at the 5% level.

The different economic indicators also have quite a different ranking in terms of their effect on expected outcomes. Overall, there is less variation across the measures in their explanatory power. Subjective poverty now has the strongest association (in terms of the parameter estimate) rather than the weakest. However, because very few people actually describe themselves as poor or very poor, the variance explained by this variable remains low. The hardship, joblessness and income support poverty variables are next, followed by the income variables and then the geographic indicators. Equivalent income has a slightly weaker relationship than income on its own – implying that children in large families perform better on social and emotional outcomes. The postcode SEIFA score actually has a slightly stronger correlation than the CD-level score.

As in the case of Table 5, the bottom half of the table shows a similar analysis undertaken for a higher disadvantage threshold. Here the strongest indicator is having two or more hardship indicators, followed by being in the bottom 30 per cent of income (which has a stronger association than equivalent or full-time income). The two geographic indicators have the weakest associations with outcomes, though again the postcode-level SEIFA has a slightly higher effect size on social-emotional outcomes than the CD-level score.

6 Associations between child outcomes and broader measures of parental economic resources

Though the literature on the best way to measure family economic resources has tended to focus on the most disadvantaged families, these measurement issues apply equally to all families. Similarly, the relationship between child outcomes and family resources is likely to apply across the full distribution of resources rather than just to the most deprived families.

Table 7 Child outcomes and general measures of family economic resources

Model	DF	**Learning domain score** Adj R^2 (%)	**Social/ emotional domain score** Adj R^2 (%)	**Physical domain score** Adj R^2 (%)
Control variables (distal and current)	31	15.3 **	9.0 **	1.7 **
Resource variables alone				
Income (+ income^2)	2	3.3 **	4.9 **	0.5 **
Equivalent income (+ ^{2})	2	3.8 **	4.9 **	0.4 **
Full-time income (+ ^{2})	2	4.7 **	4.5 **	0.3 **
How getting on (v. comfortable to v.poor)	5	1.2 **	3.5 **	1.4 **
Events due to shortage of money in last year	7	2.2 **	4.5 **	1.4 **
SEIFA of Post Code	4	3.2 **	2.9 **	0.3 **
SEIFA of CD	4	3.4 **	3.2 **	0.1
With control variables				
Income (+ income^2)	33	15.3 **	9.8 **	1.9 **
Equivalent income (+ ^{2})	33	15.3 **	9.9 **	1.9 **
Full-time income (+ ^{2})	33	15.4 **	9.5 **	1.8 **
How getting on (v. comfortable to v.poor)	36	15.3	10.3 **	2.8 **
Events due to shortage of money in last year	38	15.4	10.7 **	3.0 **
SEIFA of Post Code	35	15.3	9.4 **	1.8
SEIFA of CD	35	15.4 *	9.5 **	1.7

Source: LSAC survey, child cohort, wave 1. N=4,396. The control variables are: sex, age (months), Indigenous, non-English speaker, 1st born, age youngest child, n children, n persons, family type, mother's age at first birth, mother smoked and parents' educational attainments. The asterisks denote the significance of each set of resources variables (** = p<1%, * = p<5%).

Table 7 summarises the association between the different indicators of family economic resources and outcomes for child cohort. Because the different variables have different metrics and in some cases non-linear relationships with the outcome variables, the adjusted R^2 is used as a measure of the strength of the association between the variables. The (unadjusted) R^2 statistic can be interpreted as the proportion of the variation in the dependent variable that is ‘explained’ by the independent variables. However, this increases with the number of independent variables in the model even if the extra variables only add explanatory power because of chance fluctuations. The adjusted R^2 corrects for this, and thus can be used to compare the fit of models with different numbers of variables. For convenience, we will also use the terminology of the proportion of the variance explained to describe the adjusted R^2, even though this is only approximately correct.

The first line of the table summarises the effect of the control variables. These explain approximately 15 per cent of the variance of the learning domain score, 9 per cent of the social/emotional outcomes, and 2 per cent of the physical outcomes (the first two estimates differ slightly from those shown in Table 2 and Table 3 because of the different missing data exclusions in this table). The DF column shows that there are 31 control variables (not including the constant). For most variables the association with the physical domain scores is very low, and so discussion here focuses on the learning and social/emotional domain scores.

6.1 Income and outcomes

The second line in Table 7 shows the association between income (and income squared) and the three outcome variables. Some 3.3 per cent of the variance in learning scores is explained, and about 5 per cent of the variance in social emotional outcomes.

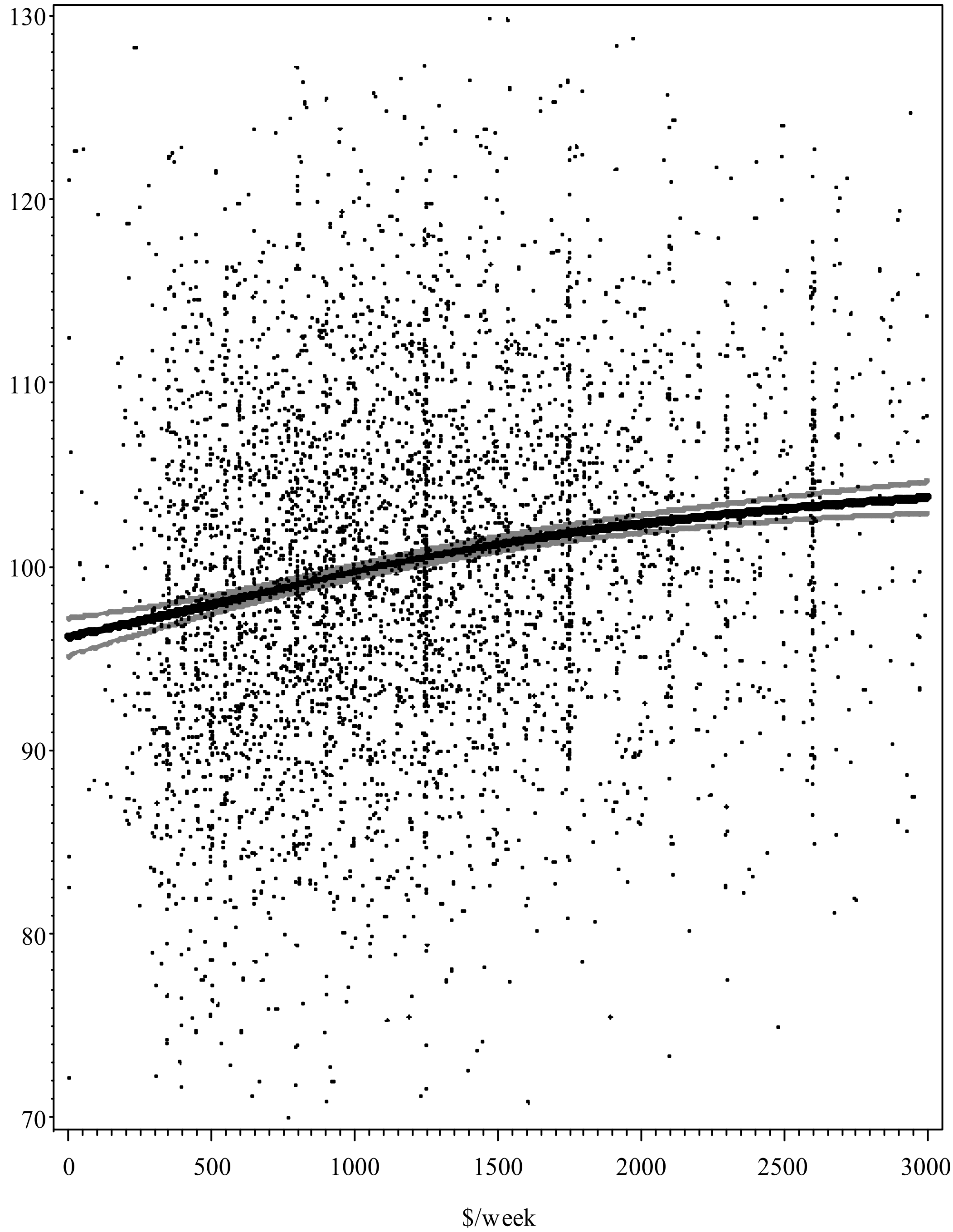
The greatest amount of variance in learning outcomes is explained by the full-time income measure, while for social/emotional outcomes, actual income or equivalent income is more strongly associated. This reflects the fact that full-time income is determined to a considerable extent by the education level of the parents, together with the fact that parental education is more strongly associated with child learning outcomes than with child outcomes in the social/emotional domain. When parental education (and other factors) is controlled for, the adjusted R^2 for learning outcomes is very similar for the three income variables (bottom panel of the table). For the social/emotional domain full-time income remains more weakly correlated.[14]

The estimates in Table 7 are based upon a quadratic relationship between the income measures and child outcomes. To test whether this is an overly restrictive assumption, Figure 2 and Figure 3 show the fitted values from non-parametric regressions as well as the raw data points and 95 per cent confidence intervals for the predicted values.[15] For the learning outcomes, the non-parametric regression produces a predicted value curve that is very close to quadratic – with the impact of an extra dollar in family income stronger at the bottom of the income distribution. For the social/emotional domain (Figure 3) there is some evidence of a weaker relationship at the very bottom of the income distribution (below \$500 per week). The scatter plots for both relationships do not show any evidence of marked heteroscedascity.

[14] One must be careful in extrapolating these results based on measured indicators to those that might occur when using other measurement methods. The full-time income variable is imputed and so may be a poorer indicator of true full-time income than is measured income is of true current income.

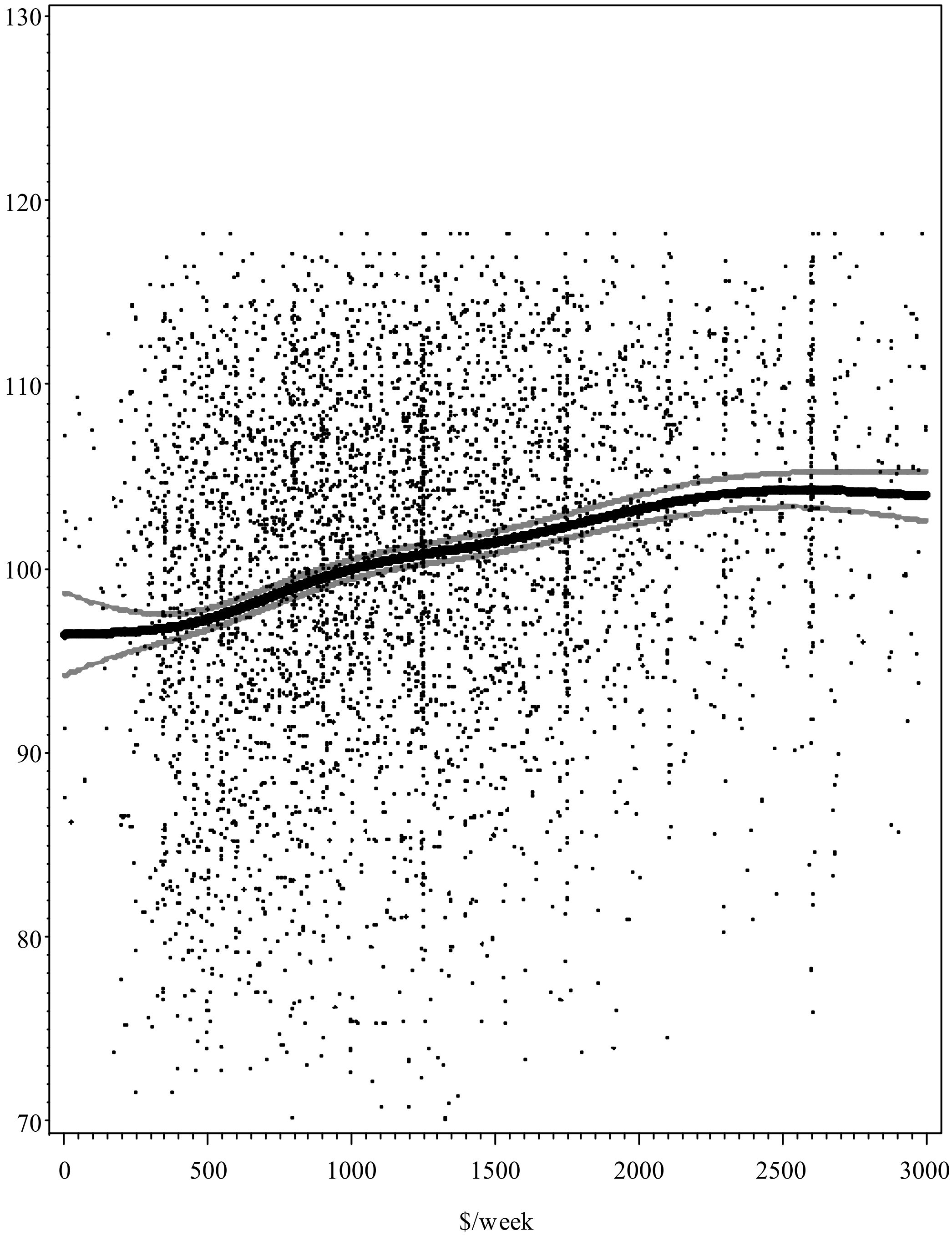
[15] The SAS thin-plate smoothing procedure was used for these calculations (tpspline). The confidence intervals assume simple random sampling (and are thus likely to be underestimates).

Figure 2 Learning domain scores by family income (non-parametric regression)



Notes: Source LSAC, child cohort, wave 1. The black line represents the predicted mean outcome score (using a non-parametric regression model). The grey lines represent the upper and lower approximate 95% confidence intervals for this.

Figure 3 Social/emotional domain scores by family income (non-parametric regression)

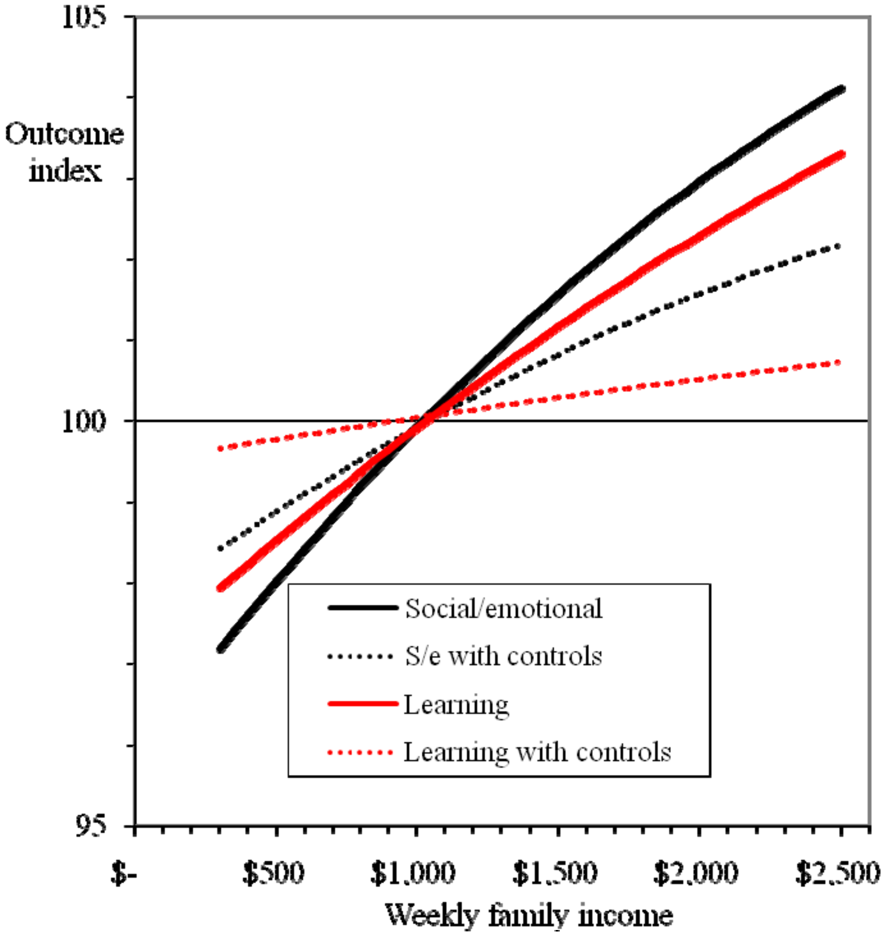


Notes: Source LSAC, child cohort, wave 1. The black line represents the predicted mean outcome score (using a non-parametric regression model). The grey lines represent the upper and lower approximate 95% confidence intervals for this.

Is the association between economic resources and child outcomes large or small? The broad scatter of points shown in Figure 2 and Figure 3 might be considered to indicate a weak relationship – the variation of outcomes across income is clearly small compared to the variation within each income level. However, this is not particularly relevant, as much of the scatter reflects the imprecision of the measurement instrument. It is not easy to measure underlying outcomes for children of this age using simple questionnaire instruments.

Figure 4 shows the predicted outcomes at different income levels corresponding to the income models shown in Table 7 (i.e. a smoothed version of the relationships in the previous two figures). Both the estimates with and without controls are shown, for both the learning and social/emotional outcome indexes. As points of reference, the 10th and 90th percentiles of family income were \$465 and \$2,375 per week respectively.

Figure 4 Predicted relationship between family income and child outcomes



Notes: Source LSAC, child cohort, wave 1.

The relationships shown imply that children from families at the 90th income percentile will have an average learning score of 4.6 points higher than those at the 10th percentile. Similarly, the gap in the social/emotional score will be 6 points. Since the standard deviation of the outcome score is set at 10, this income gap translates to about half a standard deviation – a noticeable although modest relationship.

When other variables are held constant, this gap diminishes to one point for the learning score and three points for the social/emotional score (both significantly different from zero). These effect sizes can be compared with other variables such as gender or child age. In the same regression, the effect of being female is to increase learning scores by 4.2 and social/emotional scores by 2.3. The effect of being 12 months older is 5.1 and 0.3 points respectively.

6.2 Subjective and hardship measures

The subjective living-standard question is entered here as five dummy variables. When considered in total they are only weakly associated with learning domain outcomes, and when controlling for other variables they are insignificantly associated (bottom panel). As for the poverty measures, however, a different result holds for social/emotional outcomes where this set of variables is a reasonably strong predictor on its own (though not as strong as the income variables), and one of the strongest correlates when controlling for other variables.

The seven 'shortage of money' (or 'hardship') events have a stronger association with both outcomes than does the subjective living-standard question, although they are a relatively weak predictor of learning outcomes and are not significantly associated when controlling for other variables. For the social/emotional outcomes, however, they are a relatively strong predictor, and the strongest predictor when controlling for other variables.

Inspecting the parameter estimates (not shown) for these hardship variables in the models in Table 7, the two variables 'not pay mortgage or rent on time' and 'pawned or sold something' typically had very small parameter estimates, and the 'sought assistance' variable had low estimates in some specifications. Nonetheless, an assumption that all seven hardship variables had the same effect size on child outcomes could not be rejected in any of the specifications. This provides support for the use of a simple summary scale based on the number of hardships experienced.

6.3 Geographic indicators

Two sets of four SEIFA indicators are shown in this table,[16] the first set for the postcode in which the child lives, and the second set for the CD. As noted above, because postal areas are much larger than CDs we would expect the latter to be more closely tied to the economic circumstances of the children in the study. This indeed is the case in Table 7, though the difference is not large. (Recall that in Table 6 this relationship was reversed).

For the learning outcomes, the SEIFA indicators explain less variance than the income variables do, but more than the subjective living-standard and hardship questions. For the social/emotional outcomes the SEIFA indicators generally explain less than all the other variables (though the difference from the income variables when controlling for other characteristics is small).

6.4 Summary

There are a number of points that can be made in summarising these patterns. The first is that across almost all the different economic indicators the correlation with outcomes is stronger for the social/emotional outcomes than for the learning

[16] There is no theoretical reason why the SEIFA scores should have a linear relationship with outcomes as assumed in this table. (The metric of the SEIFA scores is arbitrary). However, the addition of quadratic (and cubic) functions of each of the scores was also examined, and they were not statistically significant.

outcomes. This is also found when we focus on the bottom of the distribution (i.e. by comparing Table 5 and Table 6), and also when we control for other family characteristics (i.e. by examining the difference between the adjusted R^2 for the models in the bottom panel and the control variable model in the first row).

It is possible that this pattern reflects the different measurement approaches to these domains. The learning score includes both parent ratings and interviewer-administered items, while the social/emotional score is based solely on parent ratings. However, we would expect this to lead to poorer measurement of the social/emotional domain, but in that case the patterns would lie in the opposite direction (as is the case for the other control variables).

As noted above, this is contrary to some previous research, which has found weaker socio-economic correlations with behavioural outcomes (for older children). However, it is also interesting to reflect on the results from the intervention experiments such as the Perry Preschool Program. In surveying these results, Heckman, et al (2005) argue that these types of environmental interventions are more likely to have a long-term impact on non-cognitive skills rather than on cognitive skills. These non-cognitive skills include the types of emotional skills that the social/emotional index seeks to measure. The learning index also includes some non-cognitive skills, but is likely to be more closely correlated with cognitive skills. The fact that we find greater associations with social/emotional outcomes thus provides some support for the hypothesis that family environmental characteristics such as economic resources are likely to have a greater impact upon non-cognitive than upon learning skills.

With respect to the different measures of family economic resources, the patterns vary depending on the outcome measure. The strongest correlates with learning outcomes are the income variables, followed by the SEIFA indicators, with the subjective and hardship indicators having the weakest association. For the social/emotional domain, the last two indicators have relatively strong associations with outcomes – particularly when controlling for other demographic characteristics.

7 Income and family size

The comparison of the impact of income with that of family size is particularly interesting because of its link with the concept of equivalent income. It is normal in most poverty and inequality research to measure the economic welfare of households using equivalent income rather than actual income.

For a given household income, the average level of consumption that can be obtained by each individual in the household will be smaller as the number of people in the household increases. Because of the economies of sharing in households, however, personal consumption levels will be greater than per-capita household income. Hence researchers usually divide income by an estimate of the number of 'equivalent adults' in the household in order to obtain an estimate of the economic well-being of each household member.

The square root of the number of people in the household is perhaps the simplest equivalence scale, because it lies half-way between household income and per-capita income (income divided by the square root of the number of people in the household is equal to the geometric mean of household income and per-capita income). More sophisticated scales take account of the characteristics of the household members (such as their age). As well as being used for distributional research, these scales are also embedded in the rules used for income support programs that provide different amounts of support to families with different characteristics.

However, there is no consensus about how to estimate equivalence scales. This is because there is no clear metric available to measure individual consumption levels, as the consumption of most items in the household is shared, at least to some extent.

Can measures of children's outcomes be used as a wellbeing index to estimate an equivalence scale? For learning outcomes, outcomes are indeed lower when children have more siblings – family income (and other things) held constant (this is why the learning outcome fit is better for equivalent income than for income in Table 7). This does suggest that economic resources matter, because the average member of a larger family will have a lower level of consumption than the member of a smaller family at the same income level.

However, there are two other potential explanations for this. First, family size might have a direct effect in ways other than via their impact on consumption. For example, children might compete for the time of the parents. Second, there might be a selection effect, with parents who have large families being more likely to have lower incomes. For example, the mother might choose to have more children if her opportunity cost of time is lower.

It is probably these two factors that are responsible for the fact that social/emotional outcomes actually (slightly) *increase* with family size. Having extra children in the household might aid in the development of social capacities, and parents who are more social might prefer to have more children.

Setting these concerns to one side, Table 8 uses the relationship between learning outcomes and family size to show how much additional income would be required to offset the decrease in child learning outcomes associated with a larger family. To keep the analysis simple, the calculation here is restricted to two-parent families with 1, 2, 3 or 4 children. The first panel shows the result of a regression of learning outcomes as a linear function of income, as well as dummy variables for the number of children. The second panel shows the results of a regression where the other relevant control variables are also included.

The first column of the table shows the parameter estimates. A $1,000 per week increase in family income (about the difference between the median and the 90th percentile) leads to a 1.6 point increase in learning outcomes. Having four rather than two children in the household, however, leads to a 2.7 point decrease in learning outcomes.

Table 8 Equivalence scales for learning outcomes

Variable	Parameter estimate	Increase in income required to have same average outcome	(approx 95% confidence interval)	Required income / average income of 2 child family	(Root N equivalence scale)
No controls					
Income	0.0016				
N children = 1	-0.08	50	(-658 to 758)	103%	87%
N children = 2	0	0		100%	100%
N children = 3	-1.35	838	(407 to 1,269)	156%	112%
N children = 4	-2.73	1696	(1,048 to 2,344)	213%	122%
With controls					
Income	0.0004				
N children = 1	0.20	-125	(-2,955 to 2,705)	92%	87%
N children = 2	0	0		100%	100%
N children = 3	-0.86	534	(-1,258 to 2,326)	135%	112%
N children = 4	-1.46	909	(-1,850 to 3,669)	160%	122%

Notes: Source LSAC, child cohort, wave 1. Population: couple families with 1, 2, 3 or 4 children. Confidence interval calculated using the delta method assuming simple random sampling. The required income is the mean income (approx $1500) plus the increase in income.

The second column shows how much of an increase in income is required to offset the decrease in income associated with a larger family. This is calculated as the negative of the ratio of the parameters in the first column (i.e. $-(-2.73/0.00016) \approx 1700$). An approximate 95 per cent confidence interval for this increase is also shown. (Note that the confidence intervals are very wide for the with with-controls estimates).

The fourth column then expresses this higher income level as a fraction of the average income of a two-child family. This can thus be interpreted as an equivalence scale for the mean family.[17] As a comparison, the last column shows the equivalence scale that would be obtained by using the square root formula (with the two-child family as reference).

[17] This approach is algebraically equivalent to the Rothbarth method of estimating the costs of children. For the Rothbarth method, pure adult goods are used to estimate the cost of children on parent consumption levels. Here, child learning outcomes are used as a measure of child outcomes. See Bradbury (1994).

The striking result from this calculation is the very high cost of having many siblings. Compared to having only one sibling (i.e. a two-child family), having three siblings decreases learning outcomes so much that a more than doubling of household income is needed to offset this (for the average family). The root-n equivalence scale, on the other hand (similar to most equivalence scales in common use) assumes that family income would only need to increase by 22 per cent. Indeed, even if we assumed that every person in the household cost the same amount, and there were no sharing economies (the per-capita case), income needs would only increase by 50 per cent (6/4).[18] If we control for other variables, the impact is still very large (60% increase in income) but not statistically significant.

These high costs arise from the large impact of family size compared to the relatively small offsetting impact of income. Indeed the fact that the effect of family size is so large compared to income is strongly suggestive that factors other than the spreading of financial resources across the family are also important in large families. As noted above, these other factors could include the reduction in parental time associated with large families, or selection effects associated with the different types of parents that have large families.

The fact that social/emotional outcomes actually increase with family size can be read as implying that these direct effects of family size can more than outweigh any impact of the spreading of financial resources.

18 These differences are statistically significant if we don't control for other variables. Ie the lower bound for the 4-child family of $1048 translates to an income ratio of 170% or an increase of 70%.

8 Multiple economic indicators of outcomes

8.1 The independent impact of each economic indicator

When we consider the impact of the different economic indicators one at a time, they all have significant associations with child outcomes (and most of these are still significant when controlling for other family characteristics). Is this because these different economic indicators are highly correlated with one another, or does each of them contribute independent information on their relationship with child outcomes?

Table 9 examines this issue by testing whether these economic variables are significantly associated with child outcomes while holding constant all the other economic indicators. Because some of the variables are by definition closely related, we restrict this analysis to the variables: income, full-time income, how getting on (5 binary variables), hardship events (7), and the SEIFA index of advantage and disadvantage.

Table 9 Tests of the independent statistical significance of the socio-economic indicators

Model	DF	**Learning domain score** p	**Social/ emotional domain score** p
Resource variables alone			
Income	1	0.188	**0.031**
Full-time income	1	**0.000**	0.715
How getting on (v. comfortable to v.poor)	5	0.687	**0.001**
Events due to shortage of money in last year	7	**0.001**	**0.000**
SEIFA (advantage/disadvantage) of CD	1	**0.000**	**0.000**
With control variables			
Income	1	0.776	**0.040**
Full-time income	1	0.296	0.121
How getting on (v. comfortable to v.poor)	5	0.747	**0.004**
Events due to shortage of money in last year	7	0.222	**0.000**
SEIFA (advantage/disadvantage) of CD	1	**0.043**	**0.000**

Notes: Source LSAC, child cohort, wave 1. The table shows the probability that the observed value of the respective variable (or a larger value) could have been obtained even if there was no relationship in the population. These probabilities are based on simple random sampling formulae and are thus a slight underestimate. Where p-values are less than 0.05 the variable is significant at (approximately) the 5 per cent level – these cells are denoted in bold.

The top panel of the table tests whether each of these variables (or set of variables) is significantly associated with child outcomes while holding the other variables constant. The bottom panel also holds constant the control variables.

In the top panel, most variables are significant even when controlling for the other economic indicators. The main exception is the subjective living-standard question, which does not help to explain learning outcomes once the other economic variables are included (but does contribute to explaining social/emotional outcomes). The two income questions are also not always significant, but this is because they are by definition very similar (where both parents are working full-time they are identical).

When we also control for other family characteristics the picture changes, with only the SEIFA score being independently associated with learning outcomes. However, all variables except full income are still significantly associated with social/emotional outcomes. This different pattern probably arises because the control variables include parental education, which is more strongly correlated with learning that with social/emotional outcome measures.

8.2 Interactions between economic indicators

In the poverty measurement literature it is common to combine different indicators to obtain a composite indicator of disadvantage. This is often done in a non-additive manner. For example, people might be described as poor if they have a low income *and* also score poorly on a hardship of disadvantage index (Saunders and Bradbury, 2006). Do these indicators of disadvantage have a similar non-additive impact upon child outcomes?

Table 10 shows the results of tests for such interactions between income, full-time income and the SEIFA advantage and disadvantage index, and the other economic indicator variables. These tests are constructed in the following manner.

For the first set of tests (the first panel in the table), a regression model is estimated with child outcomes as a function of income (and income squared), SEIFA score (and squared), the set of ‘getting on’ variables, the set of hardship variables, plus interactions between income and the other variables.[19] This is done separately for the two child-outcome variables, and also with and without the other control variables included in the model. Once this model is estimated, tests are undertaken of the significance of the interaction terms. The income X SEIFA test, for example, tests for the significance of a variable comprising income and SEIFA score multiplied together. The Income X Getting On test examines the joint significance of income multiplied with each of Getting On dummy variables, and similarly for the Income X Hardship test. The body of the table shows the p values for the test that the interaction terms are zero. Values below 0.05 are conventionally defined as significant.

The second panel of the table is the same, except that here income is replaced with the full-time income variable. The third panel also has the same basic model as the first panel, but here the included interaction terms are the interaction of the SEIFA scores with the other economic variables.

[19] Quadratic terms are included for income and SEIFA scores to guard against misrepresenting non-linearities as interactions.

Table 10 Tests for interactions between income, full-time income and SEIFA scores with other economic variables

Model	Learning domain score		Social/emotional domain score	
	No controls	+ control Variables	No controls	+ control Variables
MODEL = Income (+sqr), SEIFA A/D (+sqr), Getting on, hardship + income X (SEIFA, Getting on, hardship)				
		p values		
Income X SEIFA	**0.00**	**0.00**	**0.01**	**0.02**
Income X Getting On	0.36	0.53	0.12	0.15
Income X Hardship	0.26	0.46	0.37	0.34
MODEL = FTIncome (+sqr), SEIFA A/D (+sqr), Getting on, hardship + FTIncome X (SEIFA, Getting on, hardship)				
		p values		
FTIncome X SEIFA	**0.00**	**0.00**	**0.00**	**0.00**
FTIncome X Getting On	0.75	0.87	0.14	0.07
FTIncome X Hardship	0.20	0.49	0.19	0.22
MODEL = Income (+sqr), SEIFA A/D (+sqr), Getting on, hardship + SEIFA X (Income, Getting on, hardship)				
		p values		
SEIFA X Income	**0.00**	**0.00**	**0.04**	0.09
SEIFA X Getting On	0.63	0.75	0.14	0.19
SEIFA X Hardship	0.62	0.41	0.21	0.16

Notes: Source LSAC, child cohort, wave 1. The table shows the probability that the observed value of the respective variable (or a larger value) could have been obtained even if there was no relationship in the population. These probabilities are based on simple random sampling formulae and are thus a slight underestimate. Where p-values are less than 0.05 the variable is significant at (approximately) the 5 per cent level – these cells are denoted in bold.

Most of the tests for interaction terms are non-significant apart from the interaction between income (or full-time income) and SEIFA score. This is significant at the approximate 5 per cent level for both outcome variables, with and without controls, in all cases but one.[20] The exception is the social/emotional score when including controls and with the other SEIFA interaction terms in the model (the third-last row).[21]

[20] Or possibly two cases if we take into account the fact that the tests are based on simple random sampling formula and so are slightly conservative.

[21] The top row and the third-last row both test for the significance of an interaction between SEIFA scores and income. The difference is that the top row tests for this in a model which also includes (i.e. holds constant) the interactions between income and the getting-on/hardship scores, while the third-last row holds constant the interactions between the SEIFA scores and the getting-on/hardship scores.

If interaction terms are not significant, this implies that children who come from families with better subjective living-standards or with fewer hardship indicators will tend to have better outcomes – irrespective of their income level. The effect of income is the same irrespective of the subjective living-standards or hardship.

If we are using child outcomes as our yardstick of child disadvantage, this suggests that all these indicators contribute something to our knowledge of child disadvantage, and that any composite indicator of living standards could be based on an additive index of income plus these other indicators, rather than requiring that both be present in order to indicate hardship. For example, a child from a family with a very low income but no hardship indicators would have the same outcome (on average) as a child with both a moderately low income and some hardship indicators.

The picture is however, somewhat different for the relationship between income and the SEIFA score (of advantage and disadvantage) of the child's residential location. Figure 5 and Figure 6 examine this interaction more closely for learning and social/emotional outcomes respectively. These figures show the results of bivariate non-parametric regression estimates of the relationship between income and outcomes, calculated separately for children living in each of the SEIFA advantage and disadvantage quintiles. In Figure 5, the bottom black line shows the predicted learning outcomes at different income levels for children living in the one-fifth of collectors' districts with the lowest SEIFA scores. The top, grey, line shows the impact of income for those children living in the most advantaged locations. The strength of the relationship between income and outcomes is indicated by the slope of the lines.

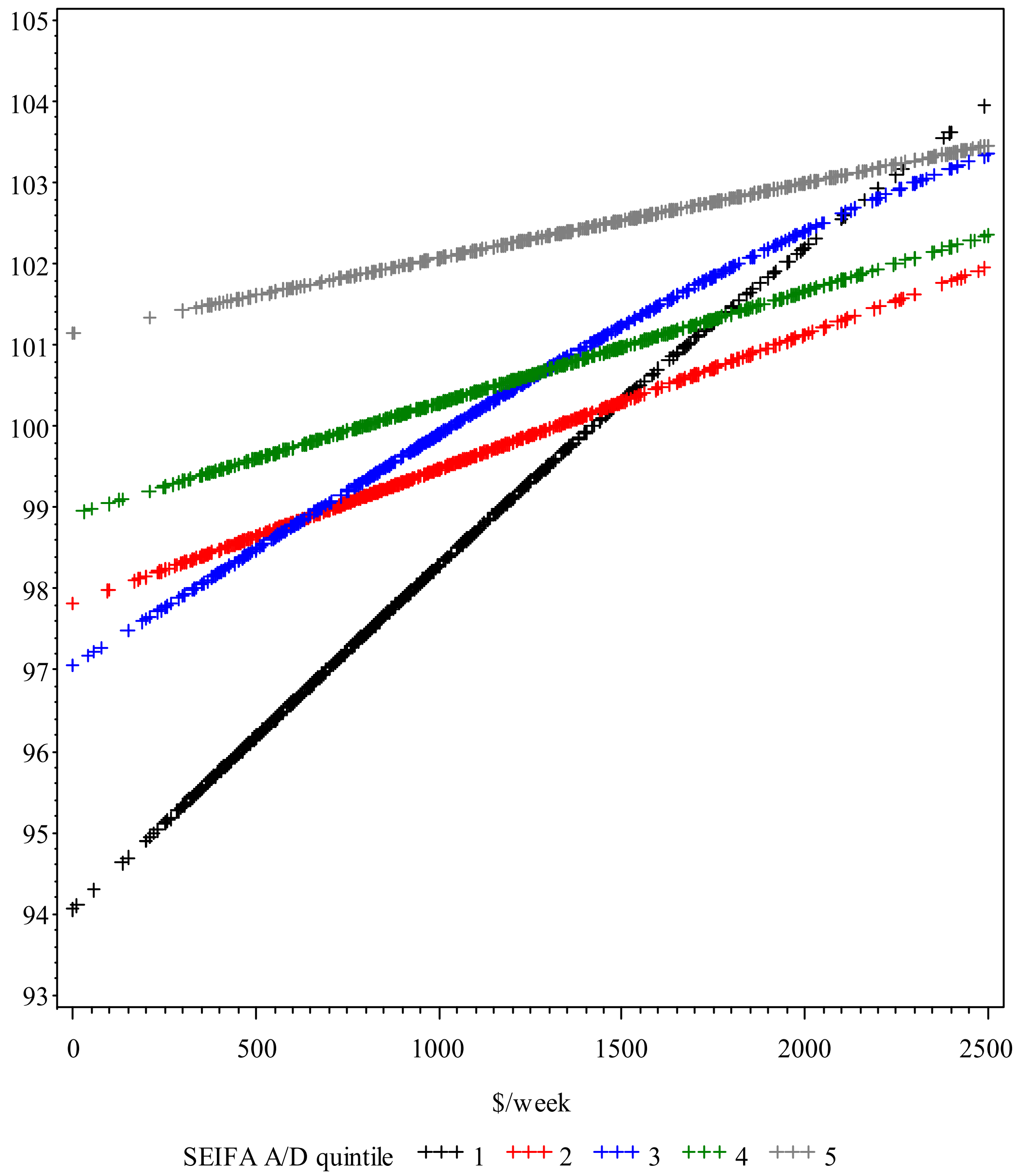
In Figure 5 the relationship between income and learning outcomes is noticeably stronger for children living in the most disadvantaged 20 per cent of regions. The slope for the middle quintile is intermediate in strength while that for other quintiles is relatively flat. The consequence of this is that for children in high-income families, the impact of region is relatively small – there is little dispersion between the lines on the RHS of the figure. For low-income families, however, there is a much greater variance of outcomes across the regions. For social/emotional outcomes (Figure 6) there is less variance across the SEIFA areas, though the slopes for the bottom two quintiles tend to be steeper than in the more advantaged regions.

Why is the impact of income on learning outcomes much stronger in the poorer regions? We can advance two possible explanations.

The first is related to how location might influence outcomes. Children in more advantaged regions have access to better quality services (childcare, health) and their peers come from more advantaged families. Thus, even children from low-income families might have good outcomes in these locations. On the other hand, if a child lives in a disadvantaged region, but the family has a high income, then they can buy their way out of disadvantage. For example, they could purchase private health care and access childcare services in other areas (e.g. near the parent's workplace).

The problem with this explanation is that it places a great deal of weight on the quality of local services and on peer effects. Are these likely to have such a large effect on the learning outcomes of children aged 4-5?

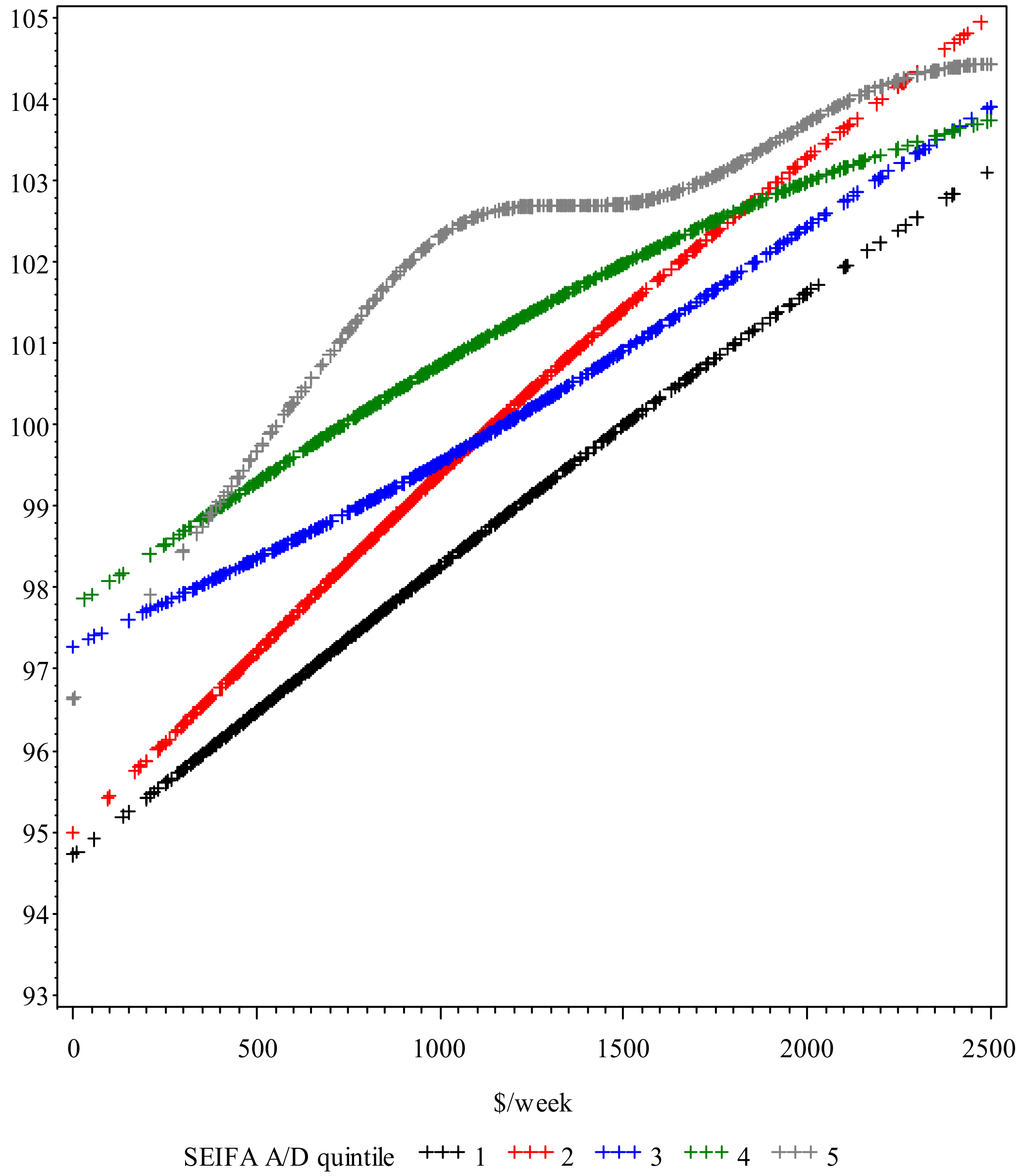
Figure 5 The relationship between learning outcomes and income by SEIFA advantage/disadvantage quintile (non-parametric regression)



Notes: Source LSAC, child cohort, wave 1. Each point represents the actual income and predicted outcome score for one child in the sample. (Note for viewers of non-colour versions of this document. At the $1000 per week income level, the lines are ranked by SEIFA quintile with the highest quintile at the top).

An alternative, and perhaps more plausible, set of explanations arises from a consideration of potential measurement problems in the income and/or SEIFA indicators. As Figure 5 shows, low-income children living in the most advantaged regions do almost as well as the high-income children in these regions. Maybe they are able to afford to live in these regions because they have access to other resources (e.g. wealth held in a business). Or possibly these families have had higher incomes in the past or expect to have higher incomes in the future. Similar types of measurement explanations can be advanced with respect to the SEIFA indicators. Low-SEIFA collectors' districts which include high- (or even middle-) income families are more likely to be socially heterogenous (e.g. in the inner regions of large cities). In this case, the SEIFA score will be a poorer explanation of the characteristics of each family. All these explanations relate to the characteristics of the parents rather than to the characteristics of the regions.

Figure 6 The relationship between social/emotional outcomes and income by SEIFA advantage/disadvantage quintile (non-parametric regression)



Notes: Source LSAC, child cohort, wave 1. (Note for viewers of non-colour versions. At the $700 per week income level, the lines are ranked by SEIFA quintile with the highest quintile at the top).

9 Further work: the development of a composite economic resource indicator

This report has examined the relationship between different indicators of family economic resources and child outcomes.[22] All the indicators considered here have substantial associations with child outcomes in at least one model specification. Even when controlling for the other indicators, most appear to contribute some additional information to the prediction of child outcomes – particularly in the social/emotional domain. Would it thus make sense to combine these different variables into a single indicator of economic resources? This section considers the different approaches that could be used to develop such an index, and discusses their suitability for a future work program.

Figure 7 and Figure 8 use path analysis notation to show two different approaches that could be used in the formulation of a composite resource indicator.[23] In both cases, it is assumed that there is an outcome variable, O, which is a function of a single resource indicator variable, R, other observed characteristics, X, and unobserved characteristics, e. R and X might be correlated (indicated by the curved line), but e is assumed to be independent of the other determinants of O. This part of the model is thus a standard regression model, except that R is not directly observed.

The 'reflective' model of Figure 7 is typical of psychometric models. It assumes that there is an unobserved underlying variable, R, which influences the observed values of the indicator variables y_1 to y_3. These indicator variables are also functions of indicator-specific error terms, which are assumed to be independent of each other. In the present context, the y variables are the observed indicators of resources such as family income and SEIFA scores (and possibly the interaction of these two). For cases where the observed indicator variables are binary (the hardship indicators) or ordinal (the getting-on questions), this model can be elaborated to make y_i, an unobserved continuous variable that influences the value of the observed variables using threshold functions.

Irrespective of the modelling strategy for the indicator variables, the key point of the reflective model is that the y variables are only associated with each other and with the outcome variable via their association with R. The resource variable thus captures that part of the y variables that tends to vary together. Moreover, since the y variables are all indicators of the same R concept, any estimates of the link between R and O should remain the same (or not be significantly different), even if a y variable is removed from the model.

---

[22] See the Executive Summary for a summary of results.

[23] See Brown (2006, Chapter 8) for more discussion of these different types of indicator models. The *reflective indicator* model is also described as an *effects indicator*, while the *formative indicator* is also described as a *composite cause* or a *cause indicator*.

Figure 7 A composite resource index: Reflective model

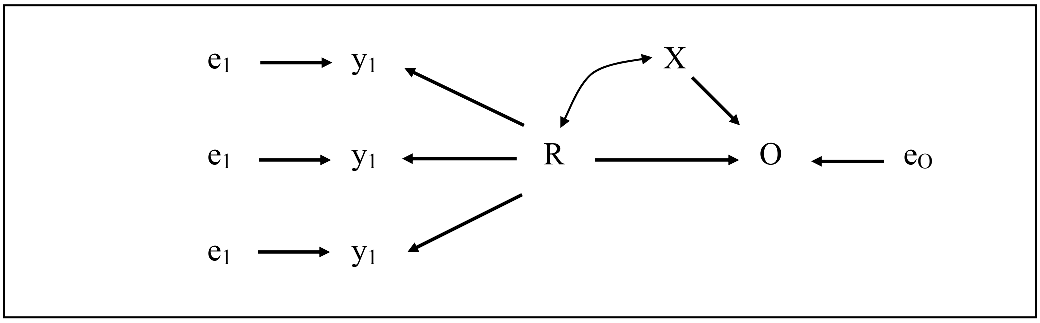
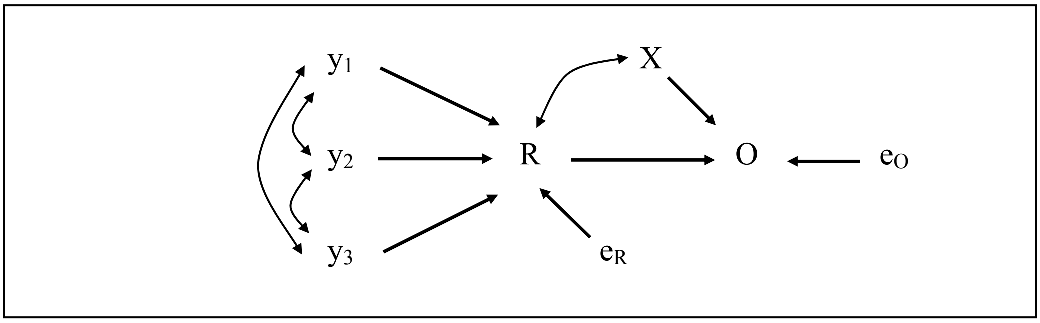


Figure 8 A resource index model: Formative model



The formative model (Figure 8) is an alternative way of summarising the impact of socio-economic factors on outcomes. In this case, the relevant concept is of a resource indicator R that is determined by the observed indicator variables. Any measurement error in the formation of this variable is captured by the error term e_R. Variants of this structure have been used to define indexes of socio-economic status which are comprised of quite different indicator variables (education, occupation, etc.). For example, the indexes of Willms and Sheilds (1996) for Canada, and of Blakemore, Gibbings and Strazdins (2006) for Australia, have this structure. For these types of indicators, the links between y and R are typically defined arbitrarily by the researchers and e_R set to zero (though in some circumstances these types of models can be estimated).[24]

Note that in this model the y variables are permitted to be freely inter-correlated (with this inter-correlation not of particular interest). This does mean that we would expect that R would vary as we added or removed y variables. For example, removing education from the Blakemore et al index would mean that it represented a different concept of socio-economic status.

The statistical identification of these two models is also quite different. The reflective model is an example of a confirmatory factor analysis model with a single unobserved

[24] The outcome indexes used in this study can also be seen as these types of formative indexes based on sub-indices which are defined as reflective indexes.

factor (R).[25] In this case, all the co-variation in the y variables is assumed to act via the R variable (because the e terms are independent). This imposes a strong constraint on the data, which can be empirically tested. If this constraint is rejected, it can be relaxed by allowing some of the error terms to co-vary, but if there aren't valid reasons for this, then the model is difficult to interpret. In the present instance, it would be theoretically plausible to allow some co-variation between the different SEIFA indexes (arising via the person's location) and between the different income variables (since they share some components). This amounts to having a multi-factor model.

The formative model as shown in Figure 8 requires additional assumptions in order to be identified. A simple one is to assume that $e_R=0$. That is, R is defined as exactly equal to a weighted sum of the component y variables. This is a very strong assumption as it is unlikely that the y variables will be measured so that they describe without error the theoretical construct that R is intended to represent. Nonetheless, it may be a sensible way to create a practical summary of the components that are considered important for outcomes.[26] It is also necessary to fix the variance of R at an arbitrary value (or fix the magnitude of one of the links between y and R).

If we exclude the e_R term, then the model of Figure 8 can be simply estimated by regression O on X and y (the X and y variables do not need to be continuous), and by re-scaling the coefficients on y to meet the arbitrary variance constraint. The predicted R variable is thus simply the sum of the y variables, with weights given by the estimates of this regression relationship (scaled by a constant).

Importantly, however, such an estimate will not be invariant to the choice of the control variables X. As new variables are added or removed from the set of control variables the OLS regression parameters will change, leading to a different predicted R.

The reflective model of Figure 7, on the other hand, is less subject to this indeterminacy as the links between y and R can be identified separately from the links on the RHS of the figure. However, with non-continuous y variables this model is more difficult to estimate than the formative model (specialist software such as Mplus is required). The single factor structure might also be rejected by the data, which might make interpretation of the results more difficult.

However, it is plausible that the variables considered in this current report are indicators of a common latent measure of family economic resources. This suggests that this reflective model will be a fruitful direction for future research. As well as developing an appropriate composite index, this would also test whether this index should be the same as or different from the two different outcome domains examined here. Given that the association between socioeconomic characteristics and child

25 Estimation is made more difficult because some variables are non-normal. However, software such as Mplus can estimate this type of problem.

26 Another identification approach is to assume that there is more than one outcome (e.g. learning and social/emotional outcomes), and that these are both determined by R *and* neither influences the other. The interpretation of type of model, however, can be difficult when the distinction between resources and outcomes is not clearly specified. Brown (2006, chapter 8) discusses similar examples of statistically equivalent models that have different interpretations.

outcomes is likely to increase as the children age, there might be merit in developing such an index using later waves of the LSAC survey.

10 References

Australian Bureau of Statistics (2004), *Mapping Australia's Progress, 2004* Cat. No. 1370.0, Canberra.

Australian Institute of Family Studies (2006) LSAC Data User Guide – Version 1.7, Melbourne

Adhikari, Pramod (2006), 'Socio-economic indexes for areas: Introduction, use and plans for 2006' Paper presented at the Census Analysis Conference, 18-19 July, Canberra.

Blakemore, Tamara, Justine Gibbings and Lyndall Strazdins (2006), 'Measuring the socio-economic position of families in HILDA and LSAC' Paper presented at the ACSPRI Social Science Methodology Conference, University of Sydney December 10-13.

Bradbury, Bruce (1994), 'Measuring the cost of children' *Australian Economic Papers* June.

Bradshaw, Jonathan (2004), 'How has the notion of social exclusion developed in the European discourse?' *Economic and Labour Relations Review* 14:168-86.

Bradshaw, Jonathan (2006), 'Child poverty and child well-being' Paper presented at the Social Policy Association Conference, 18-20 July.

Bradshaw, J. and N. Finch (2003), 'Overlaps in dimensions of poverty', *Journal of social policy* 32:513-25.

Brown, Timothy A. (2006), *Confirmatory factor analysis for applied research* The Guilford Press, New York.

Cappellari, Lorenzo and Stephen P. Jenkins (2006), 'Summarizing multiple deprivation indicators' ISER Working Paper 2006-40, University of Essex.

Cunha, Flavio, James J. Heckman, Lance Lochner and Dimitry V. Masterov (2005), 'Interpreting the evidence on life cycle skill formation' chapter prepared for the *Handbook of the Economics of Education*, E. Hanushek and F. Welch, editors, North Holland, 2005.

Dahl, Gordon B. and Lance Lochner (2005), 'The impact of family income on child achievement' National Bureau of Economic Research, Working Paper 11279.

Duncan, G. J., Yeung, W. J., Brooks-Gunn, J., & Smith, J. R. (1998). How much does childhood poverty affect the life chances of children? American Sociological Review, 63(3), 406-423.

Feinstein, Leon (2003), 'Inequality in the early cognitive development of British children in the 1970 cohort', *Economica* 70:73-97.

Fryer, Roland G. and Steven D. Levitt (2004), 'The black-white test score gap through third grade' unpublished paper, September.

Heckman, J. J., J. Stixrud, and S. Urzua (2005) 'The effects of cognitive and noncognitive abilities on labor market outcomes and social behaviour'. Paper presented at the second IZA Prize Conference (Draft version October 18, 2005). Berlin.

Hills, J., J. Le Grand, D. Piashaud eds (2002) *Understanding social exclusion* OUP, Oxford.

Kangas, O. and V-M Ritakallio (1995), 'Different methods – different results? Approaches to multidimensional poverty' in Andreβ, H-J ed *Empirical poverty research in a comparative perspective* Ashgate, Aldershot, 167-203.

Laderchi, Caterina Ruggeri, Ruhi Saith and Frances Stewart (2003), 'Does it matter that we do not agree on the definition of poverty? A comparison of four approaches' *Oxford Development Studies* 31(3),September:243-274.

Levy, D and Duncan, G J (2000) 'Using Sibling Samples to Assess the Effect of Childhood Family Income on Completed Schooling", JCPR Working Paper 168, Northwestern University/University of Chicago.

Mayer, S 1997, *What Money Can't Buy, Family Income and Children's Life Chances*, Harvard University Press, Cambridge, Massachusetts.

Morris, Pamela A., Lisa A. Gennetian and Greg J. Duncan (2005), 'Effects of welfare and employment policies on young children: New findings on policy experiments conducted in the early 1990s' *Social Policy Report* Vol XIX(II):3-17.

Nolan, B and C. T. Whelan (1996), *Resources deprivation and poverty*, Clarendon Press, Oxford.

Ringen, Stein (1988), 'Direct and indirect measures of poverty' *Journal of Social Policy* 17:351-65.

Ritakallio, Veli-Matti and Jonathan Bradshaw (2005), Child poverty in the European Union, SPRC Working paper.

Sen, Amartya (1985), *Commodities and Capabilities*, North Holland, Amsterdam.

Sanson, Ann; Sebastian Misson, Melissa Wake, Stephen R Zubrick, Sven Silburn, Sheldon Rothman and Jane Dickenson (2005), *Summarising children's wellbeing: the LSAC outcome index* LSAC Technical Paper #2.

Saunders, Peter (2004), *Lies, Damned Lies and the Senate Poverty Inquiry Report* Issue Analysis No. 46, Centre for Independent Studies, Sydney.

Saunders, Peter and Bruce Bradbury (2006), 'Monitoring trends in poverty and income distribution: Data, methodology and measurement', *Economic Record* September, 82(258):341-364.

Townsend, Peter (1979), *Poverty in the United Kingdom,* Penguin Books, Harmondsworth, UK.

Willms, J. D., & Sheilds, M. (1996). *A measure of socioeconomic status for the National Longitudinal Study of Children*. New Brunswick: Atlantic Center for Policy Research in Education, Statistics Canada.

Zubrick, Stephen R; Grant J. Smith, Jan M. Nicholson, Ann V. Sanson, Tanyana A. Jackiewicz and the LSAC Research Consortium (2006), Parenting and Families in Australia (Draft) Australian Government, Department of Family, Community Services and Indigenous Affairs: Canberra

Appendix A: Full-time income imputation

The estimation of 'full-time income' for each parent involved two steps. In stage 1, a prediction equation was estimated describing the relationship between personal characteristics and income. In stage 2, this was used to impute full-time income for those not working full-time. These two stages were undertaken separately for the mother and father of the child (if present in the household). The description below refers to the estimation for the mother (differences for the father are noted).

Stage 1: Income prediction model

This was estimated across the population of mothers who usually worked 30 or more hours per week and whose personal income was greater than $200 per week. The income cut-off excluded a small number of people with very low incomes, who might have been self-employed or working in family businesses. For fathers, the hours cut-off was 35 hours per week (the lower cut-off for mothers was used to because of the smaller sample size of mothers with long hours).

The dependent variable was log income. Income was measured with two questions (usually) asked of the mother with respect to each parent in turn "Before income tax is taken out, how much does [mother or father name] usually receive from all sources in total" and "What period does that cover?" Amounts were converted to weekly equivalents. Cases with missing data were excluded. Mother's income was missing for 11 per cent of cases and father's income for 13 per cent of cases with fathers in the household.

Since the sample of women working full-time (and with non-missing income) was only a small sub-set (17%) of the overall sample of mothers, a Heckman selection model was also estimated with age of the youngest child included as the identifying selection variable. The parameter estimates from this model were very similar and the correlation between the selection and wage equation very low ($\rho = 0.06$, se = 0.22 for mothers and $\rho = -0.11$, se = 0.33 for fathers respectively). Consequently the estimates used in this paper are derived from a simpler OLS regression of log income as a function of the predictor variables.

The predictor variables for the regression were education level (7 binary variables indicating level of school completed and post-school qualifications), occupation (for fathers, the ASCO 2-digit occupation groups; for mothers these were grouped into 16 categories because of the smaller sample size), number of children in the household (as a proxy for employment experience), and age.

The sample size for the regression was 834 (3,033 for fathers) with an R^2 of 0.41 (0.34 for the fathers). Using a log transform ensured approximately normally distributed residuals (as indicated by quantile plots). In addition to reducing the possibility of heteroskedascity, this residual pattern also convenient for the generation of multiple imputations in stage 2 (not attempted in this report).

Stage 2: Imputation

The OLS parameter estimates from stage 1 were then used to estimate predicted full-time incomes for all mothers and fathers.

For people employed (or self-employed) but not working full-time (or with missing income data), their current occupation was used in the imputation. For non-employed people we used the response to the survey questions on the last main job held for two weeks or more. In this estimation, therefore, no allowance was made for any likely decrease in wage due to absence from the workforce.

An imputed income was calculated for all people in the sample who had non-missing information on the RHS regression variables. For the estimates reported here, the full-time income is the actual income if the person is working full-time, and the imputed income if they are not working full-time (or they have missing income data).